DOCUMENT RESUME

ED 071 339 EM 010 484

AUTHOR	Rivers, LeRoy C., II
TITLE	Introduction to Psychology and Leadership. Final Report, Part I. Summary and Recommendations on a Multimedia Instructional System for Leadership, Psychology and Management.
INSTITUTION	Naval Academy, Annapolis, Md.; Westinghouse Learning Corp., Annapolis, Md.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	TR-6-18
BUREAU NO	BR-8-0448
PUB DATE	26 May 71
CONTRACT	N00600-68-C-1525
NOTE	159p.; See also EM 010 418 and EM 010 419
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Autoinstructional Aids; Communication (Thought Transfer); Curriculum Design; Curriculum Development; *Evaluation; Individualized Curriculum; *Individualized Instruction; Individual Psychology; Instructional Design; Instructional Media; Leadership; *Leadership Training; Military Training; Models; Multimedia Instruction; *Performance Specifications; Psychology; Social Psychology; Technical Reports; Test Construction; Tests

ABSTRACT

Several studies conducted in the experimental multimedia leadership management course developed by Westinghouse Learning Corporation for the United States Naval Academy are reported in this part of the final report (part two appears under EM 010 418 and EM 010 419). The research was designed to evaluate the effects of major variations in conditions of instruction involving media and presentation forms. The research plan has the distinction of being one of the first to provide a joint examination of factors in all major categories relevant to the design of an instructional system, including media, presentation factors, task requirements, and student characteristics, and in the use of an entire ongoing course system as an experimental vehicle. Supporting documents include a 12-volume set of text-workbooks (EM 010 420 through EM 010 447), coordinated test items for each content area (EM 010 451 through EM 010 464), a set of coordinated enrichment materials (EM 010 465 through EM 010 472), technical reports and papers (EM 010 473 through EM 010 483, and EM 010 487 through EM 010 500), answer and confirmation sheets to quizzes (EM 010 485), a depth core syllabus (EM 010 486), and content outlines for each area (EM 010 501 through EM 010 512). (Author/RH)

FILMED FROM BEST AVAILABLE COPY

ED 07 329

Westinghouse Learning Corporation

Contract No. N00600-68-C-1525

FINAL REPORT - PART I
SUMMARY AND RECOMMENDATIONS ON A MULTIMEDIA INSTRUCTIONAL SYSTEM FOR LEADERSHIP, PSYCHOLOGY AND MANAGEMENT

TR-6.18

May 26, 1971

EM 010 484

TR-6.18
May 23, 1971

FINAL REPORT - PART I

SUMMARY AND CONCLUSIONS

RELATING TO RESEARCH AND EVALUATION

OF A MULTIMEDIA INSTRUCTIONAL SYSTEM

Contract No. N00600-68-C-1525

May 1971

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Prepared by:

LeRoy C. Rivers, II

Approved by:

Frank B. Quirk
Project Manager

WESTINGHOUSE LEARNING CORPORATION

2083 WEST STREET

ANNAPOLIS, MARYLAND 21401

TABLE OF CONTENTS

I.	Introduction	1
II.	Background of the Research Efforts	3
III.	Research Plans	43
IV.	Research Implementation	64
V.	Summary of Spring and Fall Research Results	69
VI.	Optimization Efforts	76
VII.	Summary of Evaluation Results	107
VIII.	Some Considerations Regarding the Design and Development for Research in an On-going Instructional System	127
	References	141
	Appendix A.	144
	Appendix B.	151

TABLES

1. System Capability Matrix and Dimensions of Presentation.....12

2. Outline of Course Structure and Media.....48

3. Summary of the Research Plan For the Spring 1970 Run.....58

4. Summary of the Research Plan For the Fall 1970 Run.....60

5. Summary of the Research Plan For the Fall 1970 Run Subsidiary Questions Using Progress Check Data.....62

6. Summary of Regression Analyses Conducted.....81

7. Fall Run Student Characteristics Correlations Independent Measures.....102

8. Segments by Volume & Media Employed...109

9. Percentage of Students, by Segment, Attaining 80% or Better on Progress Checks, before Remediation, Across Three Implementations.....115

10. Segment Progress Check Performance Across Three Implementations, Mean Percentage Correct, Before Remediation.....118

11. Mean Time by Segment Across Three Implementations.....119

12. Progress Check Mean Performance (Before Remediation) Across All Materials Developed in Each Medium, 2nd and 3rd Implementation....122

13. Mean Progress Check Performance (Before Remediation) For Media by Segment for the 3rd Run.....123

FIGURES

1. Components of Educational System..... 4

2. Course Configuration and Activities Chart.....108

I. INTRODUCTION

This report describes the results of several studies conducted in the experimental multimedia Leadership Management course developed by Westinghouse Learning Corporation (WLC) for the United States Naval Academy (USNA). These studies form part of a comprehensive investigation of factors influencing student achievement intended to guide continuing improvement of the Leadership Management course, and to have wide-ranging application in the field of educational technology.

The research was designed to evaluate the effects of major variations in conditions of instruction involving media and presentation forms. Tests of a series of hypotheses were conducted with effects of experimental manipulations measured by three types of tests reflecting accomplishment of three broadly different kinds of learning tasks. The relationship of student learning in specific conditions of instruction to individual characteristics of the student was also investigated.

The WLC research plan has the distinction of being one of the first to provide a joint examination of factors in all major categories relevant to the design of an instructional system, including media, presentation forms, task requirements, and student characteristics. The WLC

plan is unique both in the number of factors investigated and in the use of an entire ongoing course system as an experimental vehicle permitting empirical findings to be extracted relevant to the influence of each factor singly and in combination. It is expected that experiments of this type, as part of a concentrated effort in educational research, may eventually result in a comprehensive understanding of the educational process, so that an instructor may choose with confidence to the most effective media and presentation forms to teach a particular type of student a particular type of task.

Research was conducted during the first full implementation of the course in the spring semester of 1970. The research was replicated with some modifications when the course was run for the second time in the fall semester of 1970. This report will provide a description of the research plans and will present a summary of the research as drawn from the results of the two implementations. An installation run of the Leadership course was conducted in the spring semester of 1971 in which the experimental designs and controls were remov removed. This report contains a section in which the design, evaluation, and findings for the installation run are discussed. The final section of this report deals with conclusions regarding the design and development for research, evaluation, and operation of an individualized multimedia instructional system.

II. BACKGROUND OF THE RESEARCH EFFORTS

In this section an informal analysis of educational systems is presented leading to the identification of major classes of variables important in the development of a comprehensive educational technology. An important distinction between presentation and media variables is then described in relation to conceptual organization of these variable classes based on Tosti and Ball's (1969) instructional design and media selection model.

Variables in Instructional Systems

Instructional systems may be analyzed as an interactive process among components of (1) designed behavioral objectives, (2) student, (3) materials designed to change student behavior toward the designed objectives, (4) media presenting materials to the student, and (5) operational organization bringing these components into articulation. The nature of such a system is illustrated in Figure 1.

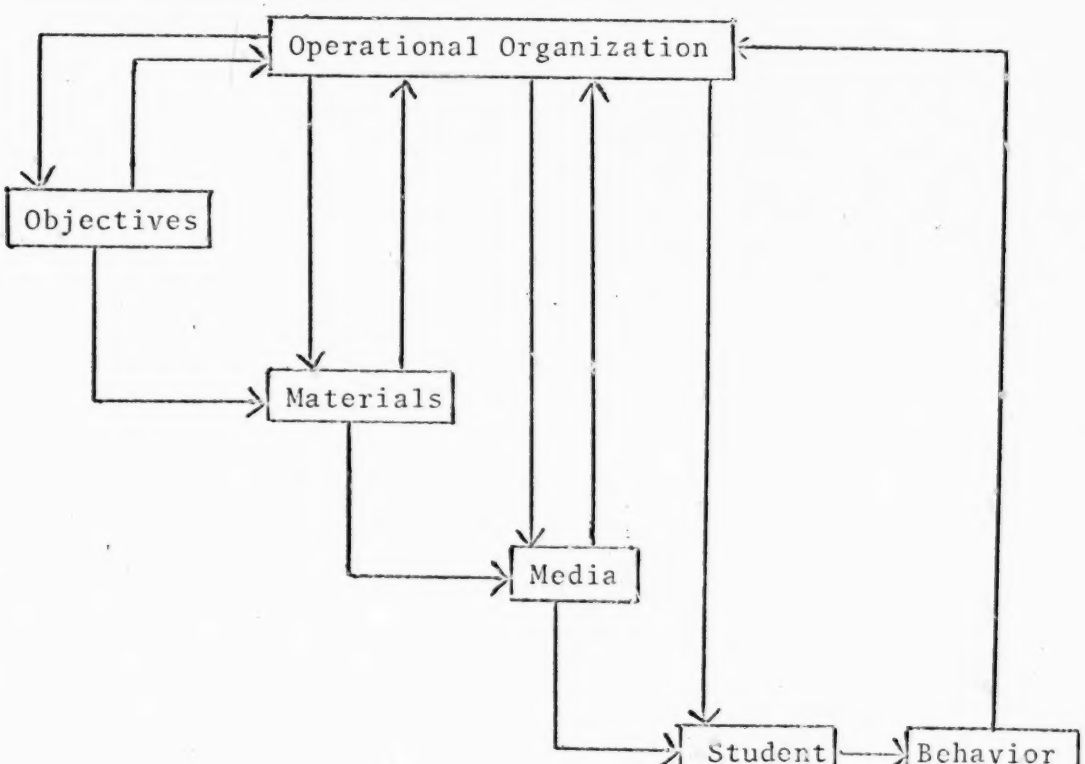


Figure 1. Components of educational system.

An analysis of this kind clarifies several points in relation to instructional systems. First, evaluation of a system rests primarily on examination of the correspondence between objectives desired and behavioral changes achieved. Other criteria, such as the desirability of objectives and cost-effectiveness are external to the system and concern the utility of the system as a component in larger social systems. Second, research on instructional systems requires the manipulation or measurement of characteristics (variables) differing among individual examples of the same component, and study of the associated modifications in the behavioral output. Clearly, evaluation of a standing instructional system may proceed, given information on objectives and behavior change, but when a discrepancy is discovered between these, efforts must return to the research domain to discover how to modify the system to eliminate the discrepancy.

Corresponding to each component of the instructional system are a large number of variables potentially important in determining the final behavior change effected. The instruction system designer is, however, not equally free to manipulate or select settings for the variables in all categories. Once objectives for

a given target population of students are established for the system, the range of variation in task variables is fixed by the objectives, and student variables are fixed by the definition of the target population. The system designer must then work with the presentation variables of the materials, media variables, and operational variables to arrange an optimally effective system for the particular types of tasks and students involved. From this point of view, presentation, media, and operational variables are of primary research interest, while task and student variables are of interest mainly in relationship to the other categories of variables.

Informal analyses of this type have provided heuristic guidelines in the development of WLC's research plan, and should prove useful in similar efforts in the future.

Presentation Variables

One of the important questions in current educational research is, "Which media will teach a given unit of instructional material most effectively?" With the introduction of so many different technological aids (including teaching machines, programed texts, television, film, cartridge tape, and computerized instruction) there is a steadily increasing variety of devices available

for use in any instructional program. The problem is to decide what medium is best for a particular purpose, and how to efficiently utilize its capabilities.

In addressing this problem, Tosti and Ball (1969) have developed a model for instructional design and media selection in which a distinction is made between medium and presentation:

> Media researchers to date have not chosen to distinguish a presentation form from the media which carry it. The new model requires that such a separation be made.
>
> The media in instructional systems carry not only the data of the instructional message, but also data on students' responses and various bits of data necessary to maintain the operating systems. It is this conglomerate of information carried by a medium which will be called the presentation.
>
> Presentation forms will be explicitly structured to communicate all data (stimulus, response, system control, student control) necessary for an efficient student-system interface. A student does not learn from the media. He learns from the presentation form. Media do little more than deliver the information to be learned in whatever presentational form previously decided upon. Some media organizations have maintained that media choice may contribute to learning efficacy because of a student's media preference characteristics or because of media dependent cues. However, the importance of these two ideas is

minimal when a separate presentation design is implemented.

To illustrate the distinction that Tosti and Ball have made, consider an instructor giving a lecture to one group of students. This lecture is videotaped and later shown to another group of students. Both groups would experience the same presentation design; only the medium would be different.

Also consider one lecture in which the instructor never answers questions, and another lecture where the instructor answers every question. In this case, the media are the same (lecture), but the presentation of the two lectures is different.

The instructional design model which Tosti and Ball have developed is essentially a taxonomy of instructional presentation variables, independent of media device, content, and external constraints. Using this taxonomy, it becomes possible to precisely describe any instructional sequence by identifying its characteristics along basic dimensions which are common to all instructional presentations. Since the specification of presentational variables is a critical consideration in educational research, the application of Tosti and Ball's model may result in a significant improvement in the quality of

studies in educational technology and in the generalizability of their findings. The presentation taxonomy may be of comparable value to other educational classification systems. As Bloom (1956) commented on the taxonomy of education objectives:

> ...(the taxonomy) is expected to be of general help to all teachers, administrators, professional specialists, and research workers who deal with curricular and evaluation problems. It is especially intended to help them discuss these problems with greater precision.

The application of the Tosti and Ball model involves the determination of a precise presentation design for each instructional objective. Media are then selected on the basis of their limitations in presenting the presentation design intact. The primary question raised by the Tosti and Ball model is whether variations of conditions of instruction in the presentation design domain are of greater or lesser importance than variation in the media domain with respect to student achievement. The implication is that if the presentation design is held constant over a unit of instruction the use of different media should not result in differential levels of achievement. Conversely variations in the presentation design with the medium constant over a unit of instruction should result in differential levels of achievement.

It would appear that a detailed behavioral analysis of individual learning events would require the specifications of a host of subsidiary variables associated with each of the Tosti and Ball presentation variables. The general approach used here was to investigate the presentation design--media controversy over larger units of instruction involving similar types of instructional objectives.

In each of the hypotheses to be tested during the Leadership Management course, the experimental treatments have been defined with reference to the six dimensions of presentation of Tosti and Ball's model. The dimensions of presentation are discussed in detail in the following section.

Dimensions of Presentation

Dimensions of presentation have been derived by a logical analysis of instructional systems (Tosti and Ball, 1969). These systems possess three basic capabilities:

a. The transmission of instructional information (stimulus capability)

b. Accepting measurable behavior of the student (response capability)

c. Changing the presentation based on the behavior of the student (management capability)

Each capability may be further differentiated in terms of two attributes: form and frequency. The result of this analysis is a 3 x 2 matrix, represented in Table 1, in which six dimensions of presentation are generated. Further study of common or possible instructional procedures reveals a number of levels or categories associated with each dimension, also listed in Table 1.

The following subsections give a detailed description of each dimension and its levels or categories. Discussions of research findings relating to the dimensions will be presented in connection with the experiments involving those dimensions.

Stimulus form (representation). This dimension is related most directly to media. It characterizes the dominant mode of sensory reception (by the student) of the instructional material, inherent in the means of representation of stimuli. There are three categories within this dimension:

a. Verbal-written -- written material, such as printed text

b. Verbal-spoken -- voice transcriptions, such as from a lecturer, videotaped lecture, or tape recorder

TABLE I

SYSTEM CAPABILITY MATRIX AND

DIMENSIONS OF PRESENTATION[a]

System Capability	Attributes of System Capability	
	Form	Frequency
Stimulus	Stimulus Representation Verbal-written Verbal-spoken Pictorial	Duration Transient-Persistant Length of time the presentation remains intact a. low b. intermediate c. high
Response	Response Demand Overt-written Overt-spoken Covert	Response-Demand Frequency Infrequent-Frequent Frequency of response required a. low or zero b. intermediate c. high
Management	Management Form Repetition Multi-level Multi-form Error-diagnostic	Management Frequency Infrequent-Frequent Frequency of decision to change presentation a. low or zero b. intermediate c. high

[a] After Tosti and Ball, 1969.

c. Pictorial -- illustrative material, such as pictures

In many presentations, two or more stimulus forms may be used simultaneously. A book may display both illustrations and prose. An educational television program conveys both a picture and lecture. Other presentations can require media-mixes such as the teacher-blackboard combination. Less common is the simultaneous employment of two variations of the same stimulus presentation, i.e., requiring the student to read and listen to the same verbal presentation.

Stimulus frequency (duration). Tosti and Ball (1969) have explained stimulus frequency as follows:

> Presentation varies on this ordinal dimension from transient to persistent, depending upon the duration of the simulus. Movies usually are conveyors of more transient presentation, and texts display relatively persistent ones. A classroom presentation by lecture is more transient than one which is delivered by the blackboard.

Transient presentations are usually instructor controlled. As in most films and lectures, the stimuli are available to the student for a fixed period of time. Persistent presentations are usually student controlled. An example would be the PI text, in which the student proceeds at his own rate and may study a unit of instructional

material for as long as he likes. Essentially, a presentation form is categorized as "transient" or "persistent" depending on the length of time the presentation stimuli remain unchanged.

Response demand. This dimension characterizes the types of behaviors which students are expected to perform in an instructional situation. The four categories within this dimension are:

a. Covert

b. Overt-written

c. Overt-vocal

d. Passive

In a PI text, the student is asked to write the answers to small units of materials. This presentation design has an overt-written response demand. A student who is asked a question in a group discussion usually answers in the overt-vocal form. The covert category describes situations where the student is asked a question, but is not required to answer with a specific, overt (observable) response. For example, the instructor giving a lecture might say, "Think of what would happen if we mixed sodium and water." The passive category describes those presentations in which questions are not overtly asked, and the student is not expected to respond

with specific overt or covert responses. Many lectures and most films are in this category. The student is only required to look and listen. However, the label "passive" should not imply that the student is doing nothing; he may be thinking intently, formulating questions about the material, or taking notes. This behavior, however, is controlled by the student, not by the instructional material. Where it is the intent of the instruction to evoke relatively specific behaviors in the learner, that presentation is categorized as either overt-written, overt-spoken, or covert.

Response demand frequency. This dimension describes how frequently the student is expected to respond (overtly or covertly) in a given period of instruction. A PI text normally has a response demand after every frame. A lecture of film may be presented with no response demands in the entire session (or module). In any medium, questions or problems may be interposed at various intervals during the instructional sequence. This presentation design would have some intermediate response demand frequency. This variable may provide a better conceptualization of what has been termed "step size" than any other.

In a temporal sequence of instruction, there are three general dimension categories:

a. High response demand frequency -- relatively frequent demand for a response in an instructional sequence, such as in programed texts where a response is required in each frame.

b. Medium response demand frequency -- relatively moderate frequency of demand for a response, such as questions which follow ten minutes of videotaped lecture.

c. Low response demand frequency -- low demand for a response, as when a "review" question follows a chapter of textual material.

Management form. Instructional management can be defined as those activities involved in the decision to assign a specific learning exercise to a given student, based on the assessment of some behavior of that student. One common example of instructional management occurs when the teacher, who discriminates that a student is having difficulty with learning a particular skill, makes the decision to assign special homework or decides to provide individual tutoring. The general logic of this activity, i.e., assessing behavior, selecting presentation, and then having the student engaging in new activity, can be extended to provide the foundation for rules employed in most new individualized instructional systems and computer-managed classroom programs.

Every instructional system involves three management elements:

a. Repertoire assessment -- appraisal of data and analysis of behavior competencies

b. Selection decision -- selection of a goal as a result of decisions based on assessment

c. Activity -- actions following from decision.

It is evident that the elements of instructional management can vary in their composition, depending on the purpose of management. Tosti and Ball (1969) isolated five purposes that may be achieved.[1] These are:

a. Need management -- to ensure students receive only those materials which they require to meet their objective.

b. Achievement management -- to ensure all students have mastered the objectives of the segment.

c. Prescription management -- to ensure a given student receives the most appropriate materials to meet the objectives in terms of his individual characteristics.

d. Motivation management -- to ensure continual student contact with the materials and to

[1]Tosti and Ball (1969) originally identified a sixth type, that of operational management. In the current presentation, this category is included among the operational system variables, since the management activities are rarely contingent on assessment of an individual student's performance.

increase student learning rate.

e. Enrichment management -- to provide for additional information relevant to objectives, but not necessary for their achievement.

The present research is primarily concerned with achievement management. Therefore, the four categories in the Management Type dimension presented below are the different procedures which may be used in management for achievement.

It frequently happens that a student is not responding to the presentation in a manner which allows him to reach the objectives. There are four strategic subclasses of management responses to such situations.

a. Repetition -- If the student fails to reach the objective, repeat the same presentation or continue through similar presentations until he does. Continuous practice is one variation of this strategy.

b. Multiform -- If the student fails to reach the objective with one presentation form, select a parallel but different form, e.g., Project PLAN (Flanagan, 1967).

c. Multilevel -- If the student fails to reach

the objectives with the presentation form, select a lower level (more expanded) form e.g., PROMOD (C'de Baca, 1968).

d. Error-Diagnostic -- If an error is made at any point within the presentation, action designed to correct that specific error is selected, e.g., intrinsic program presentation or computer assisted instruction (CAI) presentation. It is necessary when using the error-diagnostic strategy to classify errors as:

   1) input errors -- due to poor presentation design.
   2) processing errors -- due to the student's lack of the assumed appropriate repertoire on which the learning material was built, or the student's use of an inappropriate approach to the solution.
   3) output errors -- due to carelessness, poor attention, and chance error (failed to attend to a significant stimulus).

Management frequency. "This dimension is ordinal and is ordered according to the relative frequency of the decision to modify the presentation" (Tosti and Ball, 1968). The concept of decision-making in presentation design is more clearly exemplified in tutoring.

Typically, the student is directed to answer a question posed by the tutor, and a decision is made by the tutor about what he should next present, on the basis of that response. A similar instructional management form is used in PI. If an answer is incorrect, the student may be directed to any one of a number of remediation frames.

Other media may also vary in decision frequency. An instructor may ask his class a question in the middle of his lecture to see if they are understanding the material. Depending upon the students' answers, the instructor may decide to continue with the planned lecture, to review the same material, or to start a new topic. For any presentation form, the decision frequency may vary from a decision every frame to no decision at all.

In a temporal sequence of instruction, there are three general categories:

a. High management frequency -- relatively high frequency of decision to alter the presentation, based on the student's response to a question. Management frequency may be built into the instructional system, as in a text where the decision is made on the basis of a response to every frame or to remediate him on the same frame. The management frequency may also be determined extemporaneously, as when a lecturer asks a class a question; if no one

answers, the lecturer may decide to review previous content.

b. Medium management frequency -- relatively moderate frequency of decision to alter the presentation based on the student's response, such as having a quiz after a 10-minute film, and on the basis of the student's score, either repeating the film or proceeding to new material.

c. Low management frequency -- relatively low frequency of decision to alter instruction based on the student's response to a question, such as a lecturer giving a quiz after 40 minutes of lecture; basing the decision on the student's score, the instructor either assigns homework problems or does not.

It should be noted that the response-demand frequency must be equal to or more than the management frequency; decisions about a response cannot be made more frequently than one calls for that response. An example of a presentation in which response-demand frequency exceeds management frequency is the lecturer who frequently asks the class "rhetorical questions"; the lecturer does not change his presentation on the basis of the student's (covert) responses, yet he does call for those responses. In this case the response-demand frequency would be high but the management frequency would be low. (See Table 1)

Media Variables

Studies which have attempted to analyze differences between media have been largely conflicting and ambiguous. Reid and McLennan (1967), for example, reported 350 abstracts of media studies (mostly television and film); almost none of these studies found significant differences in media. Campeau (1967) selectively reviewed literature involving various comparisons among television, film, conventional lectures, programed instruction, pictoral presentations, radio and recordings, three dimensional models, and field trips. The large majority of studies reported no differences in student achievement and where differences were demonstrated (with the single exception of programed instruction) were as often in favor of one medium as another.

A number of researchers (Stolurow, 1962; Holland, 1965) have commented on the type of experimental comparison commonly attempted in studies of programed instruction. In most cases, these studies have attempted to determine the relative effectiveness of some existing instructional procedure compared to that of some new procedure or program.

This type of study has been criticized because the "existing instructional procedure, and often the new procedure or program, may be so ill-defined and poorly understood in terms of educational methodology that the results of any comparison are uninterpretable" (Ellis, 1962.)

This criticism applies to most experimental designs in which different media are compared. Media may differ in any number of ways, and be utilized in various ways. A programed instruction text, for example, presents relatively small units of material at a time, requires active responding by the student, may provide immediate feedback in terms of the correct answer, and may permit needed repetition of material. A film, on the other hand, is often viewed "passively" with large quantities of material presented in a short time, and rarely provides feedback or repetition. Even if there are differences in student achievement with these media, it is impossible to specify which elements of the instruction are responsible. In addition to the difficulty in interpreting demonstrated differences, the confounding of a large number of varied factors in "nonanalytic" comparisons also reduces the likelihood of finding any difference at all. The basis

of this latter difficulty has been clearly explained by Campeau (1966):

> .....when a single medium is used to present an entire lesson, unit or course, and achievement resulting from essentially the same presentation by an alternative medium, it is quite feasible that each medium alternately succeeds and fails in supplying the unspecified array of learning events required for the various elements of the total learning task. Whether comparisons take into account effectiveness of media or methods, or identify special characteristics of learners and media which influence learning, it is furthermore quite feasible that over the duration of a lesson, unit, or course, the net result of these alternate successes and failures, when expressed as total criterion test scores, is to conceal real differences which do exist. Hence, perhaps the great preponderance of no-difference findings in media research.

The essence of Campeau's argument is that, when presentation variables are held constant, examination of media differences at the macro-level are unlikely to succeed, since the media differences which do exist in relation to particular learning tasks and students are opposite and counterbalancing. From this point of view, micro-analysis of media variables may succeed in demonstrating media differences at the level of the individual learning event.

Tosti and Ball (1969) take an even more radical view of media, based on the implicit postulate that there are no inherent advantages of media, but only disadvantages; i.e., that a medium only makes a difference when it places some limitation on the presentation design. From this point of view, the instructional systems designer should first establish a desirable presentation design, and then select media capable of delivering that instructional presentation. Given a constant presentation design, there should be no difference in student performance resulting from delivery of the presentation through different media even at a macro-level of analysis.

Briggs (1970) has developed a model for the design of instruction in which he places emphasis on the identification of the type of learning involved in each instructional objective. Analysis of the conditions necessary to bring about each type of learning aids in determining the media to be used.
He argues that it is the responsibility of the educational specialist to define objectives and analyze learning types with sufficient precision to make obvious the necessity of particular media.

Briggs (1970) has presented a systematic method of working through the media-selection analysis, together with several examples of its application to a variety of objectives.

In the present research, WLC has compared (1) different media with the same presentation design, and (2) different presentation designs with the same medium. If significant differences are not found in the first condition, but are in the second, the generality of conclusions such as Tosti and Ball's will be supported. Such findings would serve to redirect the general research effort in media; the question "Which presentation is more effective?" may be then considered more important than the question "Which medium is more effective?"

Task Variables

Basic to the development of the multimedia Leadership Management course was an explicit statement of educational intent or educational goals for students in that course. As Mager (1968) has pointed out:

> When clearly defined goals are lacking, it is impossible to evaluate a course on program efficiently, and there is no sound basis for selecting appropriate materials, content, or instructional methods. After all, the machinist does not select a tool until he knows what operation he intends to perform. ... Too often, however, one hears teachers arguing the relative merits of textbooks or other aids of the classroom versus the laboratory, without ever specifying just what goal the aid or method is to assist in achieving. I cannot emphasize too strongly the point that an instructor will function in a fog of his own making until he knows just what he wants his students to be able to do at the end of the instruction.

Mager defines "objective" as an intent communicated by a statement describing a proposed change in a learner--a statement of what the learner is to be like when he has successfully completed a learning experience. An objective is a description of the pattern of behavior, or performance, that the learner must demonstrate. Furthermore, a statement of the objective must denote the measurable attributes observable in the learner so two independent observers

can infer correctly that the objective has been met. It is the observable and measurable character of instructional objectives which justifies the application of the term "behavioral" to such objectives.

A number of educational theorists have specified or implied that behavioral objectives involve different types of learning which may be arranged in a conceptual order from simple to complex.

Bloom (1956), for example, has written concerning his Taxonomy of Educational Objectives:

> Although it is possible to conceive of these major classes (of behavioral objectives) in several different arrangements, the present one appears to us to represent something of the hierarchical order of the different classes of objectives. As we have defined them, the objectives in one class are likely to make use of and be built on the behaviors found in the preceding classes in this list.

In presenting his Taxonomy, Bloom distinguished two broad categories of objectives: (1) knowledge, i.e., the recall of specific information, and (2) intellectual abilities or skills, including comprehension, application, analysis, synthesis, and evaluation.

Along other lines, Gagne (1965) has developed a behavior taxonomy for classifying learning tasks into eight categories:

a. Type 1 -- signal learning

b. Type 2 -- stimulus - response connections

c. Type 3 -- motor chains

d. Type 4 -- verbal associations

e. Type 5 -- multiple discriminations

f. Type 6 -- concepts

g. Type 7 -- principles

h. Type 8 -- problem solving

Gagne has argued that these learning types can be structured in a hierarchy, so that if a given instructional sequence contains more than one type, mastery of the lower-order type is prerequisite to the acquisition of the higher-order type (i.e., problem solving [Type 8] requires as prerequisites, principles [Type 7], which requires as prerequisites, concepts [Type 6], etc.).

In considering the effects of presentation design in relation to types of learning, it is important to carefully distinguish three major kinds of structure, or hierarchical organization involved in subject matter content and materials. (Briggs, 1968)

For convenience of discussion, these types of organization will be distinguished as involving content, products, and processes of learning.

The kind of organization involving content is the kind of logical arrangement of knowledge as might be

conceived by an expert in the particular discipline. As Briggs has pointed out, the organization of knowledge as an outline of the field may be entirely different than the structure needed for learning purposes. A logical content outline is often a useful means of communication between professionals who "know about" the subject-matter of the field, but is meaningless as a guide to the novice, and has no necessary relationship to the types of learning required of the novice in gaining knowledge of the field.

The organization involved in the products of learning is more related to instructional design, and involves the interrelationships among behavioral objectives which are chosen for accomplishment by the student. Analysis of these compentencies to be achieved in behavioral terms is indicative of the kind of sequencing and arrangement of elements of instruction necessary to promote efficient learning and transfer of component competencies. Questions related to this kind of hierarchical structure concern what to teach and in what order.

The process kind of organization involves the nature and sequencing of the learning events required to attain the desired competencies, i.e., how to teach what is to be taught. Questions relevant to this kind of organization primarily involve the selection and

arrangement of stimulus, response and management events designed into the materials to bring about processes resulting in a given learning product.

There is obviously an intimate relationship between the content, products, and processes of learning, but it should be clear that classifications such as Bloom's and Gagne's, refer only to the products of learning. Since classifications of products are of substantial value in the development of instruction as an aid in the analysis of content and design of materials, considerable confusion has arisen as to nature of the learning types which they identify.

It should also be pointed out that the content, products, and processes discussed above, are strictly speaking not those of learning at all, but are in fact the content, products, and processes of instruction, as conceived by an instructional designer. The structure resulting from the designer's analysis is represented in the materials developed to bring instruction to the learner, but the learner's actual behavior and modification in contact with the materials may be considerably different than that intended, even for very successful instructional materials. Since one of the goals of the behavioral analysis of instruction is to ultimately increase the

correspondence between the structure of instructional components and those of the behavior of students, the present discussion will continue to refer to "types of learning." However, some confusion may be eliminated if it is kept clearly in mind that "types of learning" refer to types of instructional product as defined in stated instructional objectives.

In the light of the distinctions discussed above, a full discussion of task variables would require a systematic analysis of content-, product- and process-related variables, and of the relationships among them. The emphasis of the present research, however, was on the relationship between products and processes. The conditions of instruction investigated were designed to affect the processes of learning in a comprehensive fashion, altering the instruction related to every objective of a segment in specified ways. The main question raised by the classifications of types of learning is whether or not the presentation variables and media have similar effects on the achievement of different types of objectives.

In order to address this question it was necessary to develop objectives and related test items measuring achievement of different types of learning. On consideration of the large-scale manipulation of conditions of instruction,

the development of tests representing rather large classes of types of learning was felt to be most in keeping with the general design of the research. The finding that particular presentation forms and media had different effects on very narrow classes of behavioral objectives would remove the advantage of instructional design at the macro-level; thus, it was desirable to determine if substantial effects of the conditions of instruction could be demonstrated using tests which include items representing several types of learning.

On the other hand, the finding of different effects with broad classes of objectives would support the procedure of segregating instruction on particular content according to the type of learning. Then the preparation of instructional units could proceed with large-scale control of presentation and media for each unit appropriate to the class of learning involved, and without major analysis at the micro-level.

Early in WLC analysis of content and objectives for the USNA Leadership Management course, it became apparent that most of the desired terminal objectives of the course could be placed at levels 7 and 8 of Gagne's (1965) hierarchy, with enabling objectives at levels 4 through 6. It was also recognized that most of the elements

involved in these objectives could be further analyzed as rules and examples in accordance with Evans, Homme, and Glaser's (1962) RULEG system. Furthermore, the objectives could be further identified according to Bloom's (1956) taxonomy as involving either knowledge of leadership and management elements, or application of those elements in realistic naval situations.

As a result of these findings, WLC developed a system of formats to be used as guidelines in the writing of specific enabling and terminal objectives. This classification scheme represents a derivation and extension of Bloom's (1956) Taxonomy, Gagne's (1965) learning types, and Evans, Homme, and Glaser's (1962) RULEG system, serving to coordinate features of each.

Test items were developed from the statement of objectives for the criterion-referenced Progress Check tests and Administrative test, which served as direct measures of achievement for purposes of evaluation. Performance on these tests thus represents learning of all four types given above, when the instructional materials are developed to explicitly teach those objectives.

The primary variables representing different classes of learning, however, were the Cumulative Posttest (CPT) items developed as special norm-

referenced research tests. These tests were developed to have approximately equal numbers of items representing acquisition of knowledge and applications. As items designed for content validity with high discriminative power, both types of items taps abilities in Bloom's other categories of comprehension, analysis, synthesis and evaluation.

Operationally, the distinction between Type I and II CPT items rests almost entirely on the presence or absence of naval situational examples in the stems or distractors. Thus there are some items which do not correspond precisely to Bloom's distinction between knowledge and applications. The use of this operational criterion of distinction, however, appeared compelling after a careful analysis of content represented in the behavioral objectives.

The specific question raised in the research is whether media and presentation variables have similar effects: (1) averaged over specific criterion behavior of all typs as indexed by the Progress Checks, and (2) on criterion-related behavior involving comprehension, analysis, synthesis, and evaluation of knowledge and the application of knowledge as measured in the CPT tests.

Student Variables

The central idea motivating research into the relationship between student variables and instructional effectiveness has been to find methods of better tailoring educational systems to the needs and abilities of individual students. Obviously, this is an area of concern intimately related to the management of instruction, but the emphasis here is on determining what student characteristics can be assessed to permit management decisions, rather than on what decisions to make given some data on the student.

Several approaches to this problem have been reviewed by Cronbach (1967). Historically, there has been much interest in selection for advancement or ability-grouping, and for this reason, research largely centered around variables predicting general academic success. On the basis of such predictors, low-ability students have been weeded out, or assigned to courses of instruction of lesser difficulty or longer duration.

An alternative approach has been to assess individual long-range goals, and areas of ability and interest, and to provide optional courses of study which appear suitable for the individual. This has been the general approach of guidance and advisement programs, providing impetus

for much research on tests in the areas of differential aptitudes and interests. More recently, this approach has been the basis of the development of large-scale computer-managed-instruction (CMI) systems, such as Plan (Brudner, 1969.) However, CMI systems are yet too new to assess their ultimate impact on individual-differences' research, since such systems have been operated primarily on the basis of a direct assessment of areas of competance, leaving the selection of goals to the teacher and student.

Only recently has major interest developed in a third approach involving the selection of a particular instructional method optimizing individual progress toward preselected goals. In the past, the selection of instructional method has been perogative of the teacher, who inevitablely modifies and utilizes methods according to his own abilities and history of success with various methods. Without standardized conditions, research on student variables predicting success under particular conditions has been difficult, if not impossible.

As Cronbach (1967) pointed out, individualized prescription of a method of instruction requires that alternative conditions of instruction designed for the same subject matter be compared in relation to student vairables to discover interactions between method and

student. That is, one should seek to discover variables for which students in one score range find one condition superior, and other students in another score range find a different condition superior.

The recent developments in the use of standardized programmed instructional materials have provided the necessary context for meaningful reserach into student-method interactions. Findings in this area have been reviewed by Stolurow and Davis (1965) and Briggs (1968).

Sufficient evidence is available to conclude that student-method interactions are quite common, if not the rule. Interestingly, variables in the areas of personality, motivation, and attitudes appear to be as important, or more important than traditional academic predictors in the findings reported thus far.

In the context of the USNA Leadership Management course, the question of general academic performance is largely moot. The students at the USNA represent a select group in terms of academic ability, and it is unlikely that variables predicting academic performance would relate to any aspect of performance in the Leadership Management course.

The purposes of research on student variables concerned the prediction of overall course

performance, and the prediction of achievement with particular media and presentation forms. Because of the number of conditions of instruction compared in the Leadership Management course, an invaluable opportunity was provided for one of the first large scale investigations of student-method interactions. To this end, a large battery of potentially predictive variables was included in the student data base.

First, the investigation attempted to identify variables predicting final course achievement. Such variables may permit the identification of students unlikely to attain satisfactory levels of course performance. Further investigation of the source of difficulty for such students may be used to find some means of remedying their deficiency. The investigation of overall performance was of general educational interest, as well, since there are few previous studies of the prediction of course achievement in the area of the social and management sciences.

Second, student variables were related to performance with particular media. Such investigations provide

information relevant to the assignment of alternate media, and on further investigation of students performing poorly with particular media, may also provide some suggestions for better accommodating particular media to the needs of individual students.

Finally, relationships between student variables and achievement with various presentation forms were investigated. The findings of these investigations may permit the utilization of the existing alternative presentations in an individually managed instructional system. In addition, some basic insights into the strengths and weaknesses of particular forms of instruction for individual students may be achieved.

Operational System Variables

A wide variety of variables must be considered when implementing an instructional system. In the traditional system the main variables dealt with are the scheduling of classes and the assignment of students and instructors to these classes. Many variables such as the length of the class periods, the grading system utilized, and the procedures for student interaction with the instructional materials are fairly well standardized.

The implementation of an individualized, multi-media instructional system necessitates a re-evaluation of many of these variables. The most obvious change involves the procedures for student interaction with the instructional materials. The degree to which students are allowed to work at their own pace must be determined. With students working at their own pace a logistical system for keeping track of the students as well as the materials must be established. Since the materials being used are not the typical "text," and since the mode of presentation used is typically not the lecture method, procedures must be established for guiding the student flow through segments of material where a variety of media are used. Once the full procedures for student interaction with the instructional materials and media are determined, consideration must be given to the personnel and facilities needed to implement those procedures. With respect to personnel, it must be determined how many students a single instructor can monitor and tutor and what additional personnel (if any) are needed to assist with record keeping and scheduling. It should also be recognized that the types of facilities as well as their arrangement will by necessity differ from those of the traditional classroom.

In a traditional classroom, time is typically held constant while performance is allowed to vary. In an individualized system, performance is held constant in the sense that all students must reach a predetermined level of performance, and the time a student spends or invests in reaching this level of performance is allowed to vary. This points out the possibility of utilizing a different set of variables to determine grades. One might, for example, base grades on the amount of time and number of attempts a student makes in achieving the desired level of performance. Final course achievement might also be based on the number of objectives achieved beyond the basic number required. If grades are indeed necessary, the type of evaluation system employed can serve as a very effective motivational device.

It should be noted here that the nature of the research involved in the course placed some artificial restriction on the operational systems variables.

III. RESEARCH PLANS

The objective of WLC's plan of research in the USNA Leadership Management course was to obtain experimental evidence relevant to the following general empirical questions:

a. Are substantial effects on student achievement produced by manipulation of presentational variables at the macrotaxonomic level as conceived by Tosti and Ball (1969)?

b. Are substantial differences in student achievement produced between different media delivering the same presentation, when measured over segments of material typical of a unit of instruction in most educational systems?

c. Are variations of conditions of instruction in the presentation design domain of greater or lesser importance than variation in the media domain?

d. Are the effects of presentation and media variables generalizable over different types of instructional objectives, or are different effects produced in relation to the acquisition and application of knowledge?

e. Are effects of particular presentation conditions and media similar for students varying according to established standarized tests of individual differences, or do the optimal conditions of instruction differ for different students?

Simultaneous accomplishment of research relevant to all of these objectives within a single ongoing course presented a number of difficulties requiring a complicated research plan. Several considerations important both to the achievement of clear-cut research findings and to the educational objectives of the USNA students in the Leadership Management course were taken into account in the development of WLC's research plan.

In performing several experiments within a single course sequence requiring repeated use of the same students it was necessary to arrange the experimental manipulation of materials and measurements so as to avoid the mutual entanglement of the effects of different experiments. Substantial variation of the level of difficulty in particular course content and test items required control to prevent obscuring of experimental effects. The small number of students available for enrollment in a developmental course required that special techniques for reducing random variation be employed to

increase the precision of the experimental comparisons, yet without interfering with the investigation of individual differences in relation to experimental variables. Finally, experimental procedures were needed which would not place an excessive burden of time and effort on the individual student, nor handicap his overall achievement through placement in ineffective learning conditions, thus leading to an undeserved reduction in course grade.

On careful consideration of all factors, a research plan was devised which substantially satisfied the criteria given above with minimal compromise among objectives. The ability of the research plan to reconcile such apparently contradictory requirements commends the WLC design approach as a model for research in ongoing courses undertaken under similar limitations.

In the outline of the research plan below, the structure of the course is described in the first subsection, with particular attention given to the cumulative posttest (CPT) unit which served as the basic research unit of instruction. The next subsection deals with medial selection and sequencing and the specific experimental designs for the spring and fall implementations of the course is discussed in the final subsection.

Course Structure

The Leadership Management course was first organized in terms of elemental blocks of content and related tests of student achievement, which were temporally sequenced without regard to research constraints. Additional elements of structure were then inserted for research purposes. This procedure insured that a basic course structure was achieved from which the research elements could easily be detached for purposes of final course packaging and implementation. The course structure may be described in terms of the four categories outlined below.

Part. The content is divided into 12 parts, corresponding to 12 chapters of the basic content outline. Each part is a formal designation of a large topic area, representing a substantial number of closely related terminal objectives relatively independent of the objectives of other parts. The objectives of any one part could be considered to be subsumed under one of the broad aims (macro-objectives) of the course. The part served primarily as an aid in fractionating the developmental work on materials.

Segment. In terms of content, a segment is a sub-collection of learning objectives within a part, which are closely related in the development of a behavioral

hierarchy of competence and in the sequencing of instructional events. A total of 59 segments were incorporated in the 12 parts of the course. The content headings of each segment are listed in Table 2 under their respective parts.

Except for revisions based on data obtained from the first implementation (Spring 1970), the content of the parts and segments remained the same for the second implementation (Fall 1970). There were some changes in medium in which some of the segments were programed. These changes from the spring 1970 to the fall 1970 run are indicated in Table 2. A more complete description of the media used in each of the two runs can be found in TR6.11, Phase II Evaluation Report and TR6.15, Phase III Evaluation Report. Changes to the research from the spring to the fall run are indicated in Table 2, and discussed in detail in the next section (Experimental Design) of this report.

Conceived operationally, the segment is the basic instructional unit in the development and production of materials, and serves as the logistical unit in implementation for purposes of scheduling and assessment of progress through the course materials. Essentially, the segment is analogous to a class period or lesson in other instructional systems, requiring 40 to 80 minutes of student time, and provides the basis for manipulation of

TABLE 2

OUTLINE OF COURSE STRUCTURE AND MEDIA

Part & Segment Number	Content Heading	Spring CPT Unit[a]	Fall CPT Unit[a]	Spring Medium[b]	Fall Medium[b]
	PART ONE: OVERVIEW OF LEADERSHIP				
1.1	Concepts of Leadership	NR	NR	ST	ST
1.2	Standards of Leadership in the Naval Service	NR	NR	F-GD	F-GD
	PART TWO: INDIVIDUAL BEHAVIOR				
2.1	Introduction to Psychology	NR	R	ST	ST
2.2	Behavior and its Observation	1	1	AT- or VT-PB	AT-PB
2.3	Learning	1	1	AT- or VT-PB	AT-PB
2.4	Factors Affecting Learning	1	1	AT- or VT-PB	AT-PB
2.5	Attention and Perception	1	1	AT- or VT-PB	AT-PB
2.6	Motivation	2	2	ST	ST
2.7	Conflict	2	2	ST	ST
2.8	Neurotic and Psychotic Reactions	2	2	ST	ST
2.9	Personality	NR	R	LAS	ST
	PART THREE: GROUP DYNAMICS				
3.1	Characteristics of Groups	3	3	AT- or VT-PB	AT-PB
3.2	The Relationship of the Leader to the Group	3	3	AT- or VT-PB	AT-PB
3.3	Group Interactions	3	3	AT- or VT-PB	AT-PB
3.4	Conformity as a Factor of Group Behavior	3	3	AT- or VT-PB	AT-PB
3.5	Relation of the Individual to the Group	NR	R	ST	ST
	PART FOUR: ACHIEVING EFFECTIVE COMMUNICATION				
4.1	Importance of Interpersonal Communication	4	4	LT	LT
4.2	Types of Communication	4	4	LT	LT
4.3	The Communication Process (Receiver and Barriers)	4	4	LT	LT
4.4	The Communication Process (Sender and Feedback)	5	5	AT-IP	AT-IP or ATS-IP
4.5	Formal Communication and Its Dimensions	5	5	AT-IP	AT-IP or ATS-IP
4.6	Informal Communication	5	5	AT-IP	AT-IP or ATS-IP
4.7	Communication Under Battle Situations	5	5	AT-IP	AT-IP or ATS-IP

Part & Segment Number	Content Heading	Spring CPT Unit[a]	Fall CPT Unit[a]	Spring Medium[b]	Fall Medium[b]
	PART FIVE: MILITARY MANAGEMENT				
5.1	Introduction to Management and the Management Process	NR	14	ST	ST
5.2	Decision Making and Creativity	NR	14	ST	ST
5.3	Objectives	NR	14	ST	ST
5.4	Planning	6	6	LT	LT
5.5	Organizing: Principles and Process	6	6	LT	LT
5.6	Organizing: Structure	6	6	LT	LT
5.7	Organizing: Charting	7	7	AT- or VT-PB	AT-PB
5.8	Directing	7	7	AT- or VT-PB	AT-PB
5.9	Controlling	7	7	AT- or VT-PB	AT-PB
5.10	Coordinating	7	7	AT- or VT-PB	AT-PB
	PART SIX: AUTHORITY AND RESPONSIBILITY				
6.1	Concept of Authority	8	8	ST	ST
6.2	Why People Accept/Resist Authority	8	8	ST	ST
6.3	Delegation of Authority; Line-Staff Relationship	8	8	ST	ST
6.4	Responsibility	NR	R	ST	ST
	PART SEVEN: LEADERSHIP BEHAVIOR AND STYLE				
7.1	Leadership Behavior	9	9	AT- or VT-PB	AT-PB
7.2	Leadership Style	9	9	AT- or VT-PB	AT-PB
7.3	Determiners of Leadership Style - The Leader	9	9	AT- or VT-PB	AT-PB
7.4	Determiners of Leadership Style - The Group and The Situation	9	9	AT- or VT-PB	AT-PB
7.5	Participative Leadership	NR	R	VT-PB	AT-PB or ATS-PB
	PART EIGHT: SENIOR-SUBORDINATE RELATIONSHIPS				
8.1	Organizational Structure & Social Distance in Senior-Subordinate Relationships	10	10	LT	LT
8.2	Officer-Enlisted Relationships	10	10	LT	LT
8.3	Assumption of Command and Formal & Informal Leader Relationships	10	10	LT	LT
8.4	Introduction to Counseling	11	11	LAS	ST
8.5	The Counseling Process	11	11	LAS	ST
8.6	Relations with Seniors and Contemporaries	11	11	LAS	ST
	PART NINE: MORALE - ESPRIT DE CORPS				
9.1	Morale	NR	R	VT-PB	AT-PB or ATS-PB

Part & Segment Number	Content Heading	Spring CPT Unit[a]	Fall CPT Unit[a]	Spring Medium[b]	Fall Medium[b]
	PART NINE: MORALE - ESPRIT DE CORPS (CON'T)				
9.2	Group Solidarity and Esprit	NR	R	VT-PB	AT-PB or ATS-PB
	PART TEN: DISCIPLINE				
10.1	Introduction to Discipline	NR	R	AT-IP	AT-IP or ATS-IP
10.2	Development and Maintenance of Discipline	NR	R	AT-IP	AT-IP or ATS-IP
	PART ELEVEN: PERSONNEL EVALUATION				
11.1	The Role of Evaluation	12	12	ST	ST
11.2	Enlisted Performance Evaluation	12	12	ST	ST
11.3	Officer Evaluation	12	12	ST	ST
	PART TWELVE: APPLIED LEADERSHIP				
12.1	Measurement of Effective Leadership	13	13	CAI	CAI or CAIS
12.2	Generally Recognized Characteristics of an Effective Leader	13	13	CAI	CAI or CAIS
12.3	Techniques of Assuming Command	13	13	CAI	CAI or CAIS
12.4	"That's an Order!"	13	13	CAI	CAI or CAIS

[a] NR refers to a nonresearch segment, thus not assigned to a CPT unit.
R refers to a research segment that did not involve a Cumulative Posttest.

[b]
ST = Syndactic (multi-level) Text
F-GD = Film, Group Discussion
AT = Audiotape
ATS = Audiotape Script
VT = Videotape
PB = Panel Book
LAS = Learning Activities Summary
LT = Linear Text
IP = Intrinsic Program
CAI = Computer Assisted Instruction
CAIS = Computer Assisted Instruction Script

the real-time parameters of the course.

At the completion of each segment, a progress check (PC) test was administered to assess the student's attainment of the terminal and enabling objectives of the segment. PC's are composed of 10 criterion-referenced items, developed directly from the behavioral statement of segment objectives.

Module. A module is a particular instructional condition used to prepare and deliver materials for a segment, identified in terms of the categories of the Tosti and Ball (1969) model. Several parallel modules were prepared in each segment utilized for research purposes, representing variations specified by the experimental designs. The different modules of a segment are distinguishable from one another by differences in presentation design and/or media, although the content is the same. Specifications of the modules for each segment are outlined in later sections of the paper giving the design of each experiment.

Cumulative Posttest unit. The cumulative posttest (CPT) unit is a group of three or four adjacent segments within a part. In the Spring 1970 run, there were 13 CPT units involving 45 of the 59 segments of the course, as listed in Table 2. The primary criteria for grouping segments into CPT units were that the segments dealt with similar types of content and objectives, and that the instructional sequences relating to particular concepts which were initiated in the unit would also terminate in

the same unit. All segments in a CPT unit were developed in the same medium and with the same variations in instructional conditions between modules.

In the second implementation (Fall 1970), there were 14 CPT units involving 48 of the 59 segments. In addition, research using progress checks as the dependent measure was conducted involving 9 additional segments.

The CPT unit is the fundamental unit of instruction for research purposes, providing the framework on which the experimental designs were constructed. The students were divided into groups assigned to different modules in the CPT unit. A student in any one group would thus encounter the same experimental conditions in progressing through the three segments of the unit, and would take three PC's, one after completing his module of each segment. After completing the segments and PC's all students then take the CPT, a test administered to assess overall achievement level under the experimental conditions represented in the CPT unit.

Performance on the CPT was the primary dependent measure for research purposes. Each CPT was composed of 10 multiple-choice items for each segment in the unit, so that CPT's for 3 segment units had 30 items, and CPT's for 4 segment units had 40 items. There were approximately equal numbers of two types of items: Type I, representing

acquisition of knowledge of the concepts and principles in the unit, and Type II, representing application of those concepts and principles in the unit in relation to realistic examples of leadership situations.

CPT items were designed to have content validity in relation to the objectives of the unit, but unlike the PC items, also to have high difficulty and discrimination power. The CPT tests thus provided norm-referenced rather than criterion-referenced measures of achievement level. Many items were designed to measure the ability to integrate behaviors from different segments in the unit. An effort was made, however, to maintain an equitable representation of content from the several segments of the unit.

Following completion of the CPT, each student is given remediation on segments where his PC test performance is below 80%. The remediation consists of repetition of the same instructional materials previously used with the segment, or materials of an alternative module thought to be more effective. On completing remediation, the student repeats the PC's for those segments and then proceeds to the next segment.

Media Selection and Sequencing

Selection of the media for the instructional design of the USNA Leadership Management course was predicated on the requirements of the experimental designs, capacities needed for delivery of instructional presentations, and the diversity and flexibility expected of an individualized multimedia course. The media selected permitted precise experimental control and planned variation in dimensions of stimulus representation, duration, response form, response demand frequency, and management decisions. Within limitations of existing facilities at USNA, media were selected which could be used in individually paced instruction without undue logistic difficulties, and with sufficient variety of instructional technique to maintain a consistent level

of student motivation. Some media were selected for their novel appeal, while the experimental manipulations of presentation design provided variety in the utilization of more traditional media forms.

Media placement and sequencing was limited to some extent by the number of segments required for each CPT unit and the number of segments in each part. Within these limitations, media were assigned to ensure perception of a sense of media variety, and to provide persistent media in segments with the most complex concepts.

The final media assignment to segments was carried out so as to provide a sufficient number of CPT units in the same media to accommodate the designed experiments, and to keep the CPT units of a given experiment widely separated in the course.

The purpose of having widely separated units in the same experiment with units assigned to other experiments intervening was to insure that any carryover effects of experimental conditions which remained after remediation were not carried systematically into the treatment conditions of the same experiment. Such carryover effects were, instead, randomized among the treatment conditions of different experiments. Thus, while the carryover effects might produce some increase in

variability and loss of precision in comparison of conditions of instruction, they were not allowed to systematically bias any comparison of conditions. Wide separation of CPT units in the same experiment also could be expected to systematically reduce problems of sequential correlation often associated with experiments of the repeated measure type.

The purpose of assigning the same medium to the CPT units of a particular experiment was the same as that given for holding the medium constant in the CPT unit, i.e., to permit the presentation variables to be manipulated in the same way and to produce similar effects in each segment of a given experiment. Since the presentation variables might possible have different effects in different media (media-presentation interactions), such variation in effects was avoided in all but Experiment III, where the interactions were of direct interest. Otherwise, the linear models for the statistical analyses would have been based on erroneous assumptions, and the interpretation of results would be somewhat difficult.

Experimental Design

Spring 1970. Table 3 presents a summary of the research plan for the spring 1970 run of the course. Experiment I involved 16 segments in which two variables were manipulated. Each student worked through two consecutive CPT units in videotape and two consecutive CPT units in audiotape. Half of the students had high RDF presentations and half had low RFD presentations. Experiment II involved nine segments in which two variables were being manipulated with the medium of linear text being used in all segments. Each of the students worked with each of the three types of response demand (i.e., covert in CPT 4, over-spoken in CPT 6 and overt-selected in CPT 10)

TABLE 3

SUMMARY OF THE RESEARCH PLAN
For the Spring 1970 Run

EXPERIMENT	CPT & SEGMENT	MEDIA	VARIABLES	CONDITIONS
				A B
I	CPT 1 2.2-2.5 CPT 3 3.1-3.4 CPT 7 5.7-5.10 CPT 9 7.1-7.4	Taped Lecture, Audio & Video (with Panel Book)	A. Response Demand Frequency (RDF) B. Media (Audiotape vs Videotape)	1. High RDF-Videotape 2. High RDF-Audiotape 3. Low RDF-Videotape 4. Low RDF-Audiotape
				A B
II	CPT 4 4.1-4.3 CPT 6 5.4-5.6 CPT 10 8.1-8.3	Linear Text	A. Response Demand Frequency (RDF) B. Form of Response Demand (RD)	1. High RDF-Overt selected RD 2. High RDF-Overt spoken RD 3. High RDF-Covert RD 4. Low RDF-Overt selected RD 5. Low RDF-Overt spoken RD 6. Low RDF-Covert RD
				A B
III	CPT 5 4.4-4.7	Audiotape/Intrinsically Programmed Booklet (AT/IP)	A. Response Demand Frequency (RDF) B. Management Frequency	1. High RDF-High MF 2. High RDF-Medium MF 3. High RDF-Low MF 4. Low RDF-Low MF
	CPT 13 12.1-12.4	Computer Assisted Instruction (CAI)		These conditions apply for both AT/IP & CAI.
IV	CPT 2 2.6-2.8 CPT 8 6.1-6.3 CPT 12 11.1-11.3	Syndactic Text	Remediation method	1. High RDF Remediation 2. Low RDF Remediation 3. No Remediation
V	CPT 11 8.4-8.6	Learning Activity Summary (LAS)	Peer Interaction	1. Peer Interaction 2. No Peer Interaction

in different sequences. Half of the students were given high RDF presentations across all segments and half were given low RDF presentations. Experiment III covered eight segments in which two variables were manipulated. Each student was presented only one of the four conditions or modules in CPT 5 and 13 and it was the same one in both CPT units. As indicated in Table 3, the medium used in CPT unit 5 was audiotape/intriniscally programed booklet, and in CPT unit 13 the medium was Computer-Assisted Instruction. Experiment IV involved 9 segments in which the remediation method was varied. Each student worked through one CPT unit in each of the three conditions. Finally, Experiment V covered three segments in which peer interaction as opposed to individual study was investigated.

The reader should refer to TR6.12a Phase II Research Findings, Part I: Conditions of Instruction for a detailed discussion of these experiments and their outcomes.

Fall 1970. The changes made to the research plan for the Fall run are indicated in Tables 4 and 5. In Experiment I the use of videotape was dropped. The new variables that were manipulated were the form of the response (RD) and provision of confirmation (knowledge of results). In Experiment II the overt-spoken RD type was replaced with one that required the midshipmen to underline their response. In the Spring run, performance where Computer-Assisted Instruction (CAI) was used was exceptionally high and there

TABLE 4

SUMMARY OF THE RESEARCH PLAN
For the Fall 1970 Run

EXPERIMENT	CPT & SEGMENT	MEDIA	VARIABLES	CONDITIONS
I	CPT 1 2.2-2.5 CPT 3 3.1-3.4 CPT 7 5.7-5.10 CPT 9 7.1-7.4	Audiotape (with Panel Book)	A. Response Demand Frequency (RDF) B. Form of Response Demand (RD) C. Confirmation	A B 1. High RDF-Confirmation 2. High RDF-No Confirmation 3. Low RDF-Confirmation 4. Low RDF-No Confirmation These conditions apply with both overt and covert responding.
II	CPT 4 4.1-4.3 CPT 6 5.4-5.6 CPT 10 8.1-8.3	Linear Text	A. Response Demand Frequency (RDF) B. Form of Response Demand (RD)	A B 1. High RDF-Overt selected RD 2. High RDF-Overt underline 3. High RDF-Covert RD 4. Low RDF-Overt selected RD 5. Low RDF-Overt underline 6. Low RDF-Covert RD
III	CPT 5 4.4-4.7 CPT 13 12.1-12.4	Audiotape/Intrinsically Programed Booklet (AT/IP) Computer Assisted Instruction (CAI)	A. Response Demand Frequency (RDF) B. Management Frequency (MF) C. Media (Script vs Audiovisual)	A B 1. High RDF-High MF 2. High RDF- Medium MF 3. High RDF-Low MF 4. Low RDF-Low MF 5. High RDF-High MF (script) These conditions apply for both AT/IP & CAI.
IV	CPT 2 2.6-2.8 CPT 8 6.1-6.3 CPT 12 11.1-11.3	Syndactic Text	Remediation method	1. High RDF Remediation 2. Low RDF Remediation 3. No Remediation
V	CPT 11 8.4-8.6 CPT 14 5.1-5.3	Syndactic Text	Advance Organizer (Content map)	1. Content map 2. No content map

was some question whether it was due to the medium itself or to the content programed for it. Therefore, an additional condition was added to Experiment III. This condition involved the use of a script version of the CAI and audiotape/IP presentations. In Experiment IV the conditions remained the same, but the students were forced through the remediation sequence which they encountered. Experiment V saw the greatest change. The LAS units were re-written in Syndactic text format, the use of the content map was investigated, and another CPT unit (involving 3 additional segments) was added.

Some questions were added to the overall research plan that involved the use of the progress check data rather than the CPT data. These subsidiary questions involved the use of nine more segments in the research effort as is shown in Table 5. In the first analysis the variables of advance organizers and the use of a special revealed answer form were manipulated. Each student saw each of the four conditions. The question had been raised whether a paper version would be as effective as the hardware bound versions (i.e., CAI and audiotape). Therefore, in segment 7.5, in addition to the use of the advance organizer, the effectiveness of the audiotape script as opposed to the tape itself was studied. The third subsidiary analysis looked at the same variables as the second but it included the use of the type with the intrinisically programed booklet as well as the panelbook.

TABLE 5

SUMMARY OF THE RESEARCH PLAN
For the Fall 1970 Run
Subsidiary Questions Using Progress Check Data

	Segment	Media	Variables	Conditions
1.	2.1 2.9 3.5 6.4	Syndactic Text	A. Advance Organizer (Content map) B. Special Answer Form	1. Answer form - map 2. Answer form - no map 3. No answer form - map 4. No answer form - no map
2.	7.5	Audiotape/Panel Book Audiotape Script/Panel Book	A. Advance Organizer (Content map) B. Medium (Audio and script)	1. Audiotape - map 2. Audiotape - no map 3. Audiotape Script - map 4. Audiotape Script - no map
3.	9.1 9.2 10.1 10.2	Audiotape/Panel Book Audiotape Script/Panel Book Audiotape/Intrinsically Programed Booklet Audiotape Script/ Intrinsically Programed Booklet	A. Advance Organizer (Content map) B. Medium (audio and script)	1. Audiotape Script - map 2. Audiotape - map 3. Audiotape Script - no map 4. Audiotape - no map (The above conditions apply for the Panel Books [9.1-9.2] and the Intrinsic Booklets [10.1-10.2]

Each student saw each of the four basic conditions, two with the panelbook and two with the intrinsic booklet.

The final difference between the two implementations dealt with the dependent measures used. In the spring run, the data were analyzed with respect to type I and type II CPT scores as well as Total CPT scores. In addition, the analyses were conducted using progress check performance as a dependent measure. In the fall run only total CPT scores were used in the major experiments. Progress check data were used, however, in the subsidiary analyses.

The reader should refer to TR-6.16 Phase III Research Findings for a detailed discussion of these experiments and their outcomes.

IV. RESEARCH IMPLEMENTATION

This chapter discusses the two research runs. The installation run in the spring of 1971 is discussed in Chapter VII.

Students

Forty-four third classmen (sophmores) from the United States Naval Academy were enrolled in the Lead Leadership, Psychology and Management course in both the spring and fall runs.

Before commencing work on the course materials, each student was randomly assigned to a track. This student track indicated the precise module of materials a student would use in each of the 59 segments of the course.

Materials

The basic course structure was discussed in Chapter III, and Table 2 presented an outline of the course structure and the media used. The segments listed in Table 2 are core segments. That is, they are required segments which include all of the information pertinent to the attainment of the requisite behavioral objectives. In addition to core segments there were depth core and enrichment segments. Depth core segments were associated with one or more segments and were directed toward amplifying the learning objectives of those segments. Depth core segments included in the

second implementation were film, group discussions, and classroom lectures by the USNA instructor. Unlike core segments, depth core were scheduled by the instructor with respect to time and place. Student attendance was required. Enrichment segments were related to but not essential to the mastery of terminal objectives. They were optional to students who desired more information than that presented in core segments.

Tests

Four different tests were used throughout the course. They were the administrative pretest and post-test, the progress check, the cumulative posttest (CPT), and the USNA examination.

The administrative pre and posttest was an 80 point criterion referenced test composed of items representatively sampled from the objective-test item pool. There was at least one administrative test item for each segment of the course. The pretest was given at the beginning of the course, and the posttest was given as part of the final examination.

The progress checks were a criterion referenced test of approximately ten items. They were given at the end of each segment.

The cumulative posttest (CPT) was a norm referenced research test composed of positively

discriminating content-related test items. Each CPT was composed of ten items for each segment in the unit. Cumulative posttests were given at the end of each experimental unit.

USNA examinations were a combination of criterion referenced test items selected from the objective-test item pool and items developed by the USNA on-site instructor. These were the only tests in the course which were used to determine the midshipmen's grades.

Procedures

The second implementation of the course was conducted in the Fall of 1970. The course was administered by the USNA on-site instructor, the WLC on-site instructor, and a course administrator. Detailed procedures used in implementing the course are given in the Instructor's Guide (TR6.6).

The instructor's basic responsibilities were tutoring students needing remediation, leading group discussions, scheduling and administering depth core segments, scheduling and administering examinations, and determining grades.

The course administrator developed and supervised the logistical procedures of the course, controlled dissemination and collection of all core materials, remediation prescription forms, module questionnaires, progress checks, and cumulative posttests (CPTs). The

course administrator also scored progress checks and CPTs and forwarded data to WLC's computer center.

Students were routed through the course according to procedures outlined in the Student Guide (TR6.5). In brief, students worked through core segments of the course at their own speed. They were allowed to check out materials and study them whenever and wherever they wished. All students were given identical material when they studied a non-research segment; i.e., they were instructed by the same form of presentation. For research segments, they studied by the particular module (form of presentation) to which they were assigned. Students were randomly assigned to modules at the beginning of the course. Each student received his own routing schedule which included not only the sequence of segments he must study but also the schedule for remediation, research tests, and USNA examinations.

Students worked through non-research material by studying a segment, taking a progress check, remediating (if necessary), and then retaking the progress check. The requirement for remediation was based on failure to attain 80% of the objectives as measured by the progress check. If the student failed to meet the 80% criterion on his first try, he was given a remediation prescription form which directed him to specific points in the materials which related to the objectives failed.

If the student failed to meet the 80% criterion following remediation, he reported to the on-site instructor for tutoring.

Students worked through research segments in the same manner as non-research segments except that they did not remediate until after they had completed the entire research unit and taken the cumulative posttest.

Facilities

For the implementation of the course, WLC was provided three classrooms at the Naval Academy. One room which was designated as the administrative office contained desks for the administrative staff and storage space for half of the course materials (including tapes, printed material, tests, forms, and computer cards). The administrative room was used as the site for administrative conferences, for student tutoring, and for distribution and collection of all material.

The second room was used as the principal instruction room. It contained 15 student carrels.

The third room, used as a multi-purpose room, had three carrels to handle overflow from the instructional room. In addition, there were 30 student writing desks which were used during depth core lectures, films, group discussions, and testing.

V. SUMMARY OF SPRING AND FALL RESEARCH RESULTS

This research effort was directed at answering the central question of whether variations of conditions of instruction in the presentation design domains as posited by Tosti and Ball (1969) are of greater importance than variations in the media domain. Tosti and Ball's position has not received strong experimental support from this series of studies. Indeed, in Experiment I of the spring run where the presentation design was held constant, a significant difference was found between media. Although this does not necessarily refute the basic assumptions of the Tosti and Ball model, it does indicate that other variables relating to the production of instruction particularly via transient media (videotapes and audiotapes) should be considered. On the other hand, one can see many cases where no significant differences were found when the medium was constant and the presentation design was varied.

In investigating the question of whether substantial effects of student achievement could be produced by manipulation of presentation variables over large segments of instruction, Experiment II of the spring run provided the most conclusive evidence. It was found that with linear text, the high response demand frequency condition was

consistently superior to the low response demand frequency condition. A facilitative effect does appear to be produced by the insertion of a significant number of questions in the instructional material. Further investigations need to be conducted with respect to this variable as implemented with transient media (see Experiment I spring and fall runs).

Although major differences were not found with respect to the form of response required of the student, the trend in Experiment II of the spring run was that the overt-selected response condition was slightly better than either overt-spoken or covert. This finding is generally in concert with the major body of prior research that indicates that the form of response utilized within the instructional materials should be similar to that required of the student on the tests of achievement on those materials. In Experiments I and II of the fall implementation, no significant differences were found with respect to the main effect of the form of the response demand (RD). However, if one has a high RDF presentation it was felt that the most effective RD might be overt-selected which is in concert with the spring findings.

Perhaps the least conclusive finding related to manipulations of management frequency (MF) in conjunction with variations in response demand frequency (RDF) as indicated in Experiment III of the spring run. It appeared that the most efficient condition was the use of moderate levels of management frequency (MF) in conjunction with high response demand frequency (RDF). However, an intervening variable may have been the fact that knowledge of results was not provided in those cases where management was not manipulated. This may have produced variations in the experimental conditions that were not due to the major variables investigated. In the fall run, with confirmation added, the results showed no significant differences among the combinations of RDF and MF nor between a script version of the high RDF-high MF condition and the automated (tape or CAI) version.

In Experiment IV of both the spring and fall runs the type of remediation provided after the syndactic text summaries had no effect on the students performance. They performed equally well. Indeed it was the case that the students generally performed so well on the syndactic text summaries that a precise test of the remediation conditions in this medium was difficult to obtain.

Indeed the preponderance of findings were "no significant differences"! In Experiment V (fall run) no significant differences were found with respect to the use of "content maps" as advance organizers. This was replicated in the subsidiary analyses of the fall run using progress checks as the dependent measure. In these same subsidiary analyses performance was almost identical where scripts of tapes were compared with the use of the tapes themselves. With respect to both media and presentation variables, there were no consistently significant differences with respect to the type of learning or task required of the student as seen in the first implementation. That is, the findings with relation to norm-referenced CPT performance were generally consistent with the findings in relation to progress check performance.

Comparison among media with the same kind of control and precision devoted to comparison of presentation variables in the present studies may well have demonstrated numberous reliable differences among media, as was the finding in Experiment I of the spring run. However, these results do support the general conclusion that differences among media, if they exist, are no more substantial nor important than variation in difficulty among

units in the same media, resulting from variation in content programing skill, or test items. Thus it would seem that the problem of media selection should deserve no more attention from the educational technologist than he is willing to devote to these other sources of variation in student achievement.

General Conclusions. With such a large number of analyses showing no significant differences one has the choice of accepting the fact that there is a high probability that no real differences exist or that the experimental designs and analyses are in someway at fault.

Much of the prior research on media and various presentation variables involved instructional materials of a short-term nature which usually covered one main content area, and were frequently conducted in an experimental environment. It should be noted that this research effort involved investigations over a semester's worth of instruction in a required course for a significant portion of the midshipmen at the U. S. Naval Academy. The content of the instructional materials covered topics in many disciplines (see Table 2). In addition, many of the experiments were replicated in a second implementation of the course.

The experimental designs were sound and employed tight control to ensure that the research data could be gathered without hindering the learner's opportunity to learn. The part and segment results indicated significant differences as small as 2.5 to 3.0 percent. Although they were statistically significant they obviously were not educationally significant. Therefore, it was felt that it was unlikely that the lack of findings of effects of the primary experimental variables could be attributed to a lack of precision in the experimental designs and analyses.

The above discussion does not mean however, that it has been established as fact that no real differences exist with respect to media and presentation variables. What it does indicate is that with general prescriptions to the programers (i.e., for high RDF write a question for every frame; for medium MF provide response sensitive feedback every other frame, etc.) and precise guidelines, based on a sytematic approach, for the revision of the instructional materials, almost any medium with almost any presentation design characteristics can be programed so that the learners can reach a pre-set level of acceptable performance. However, certain subject matter programed in some media, with certain presentation

characteristics, will demand more effort and cost to show the desired level of performance. Some may consider cost the most critical factor and opt for slightly less effective student performance. Others may well consider effective student learning to be most critical, at whatever the cost.

The topics for future research that would appear to have high payoff for instructional technology would be the investigation of optimum revision procedures and the delineation of the variables important in developing effective programing skill. Indeed, the programer needs to be an independent variable in future media research just as the teacher and test administrator have been recognized as independent variables in other research. Even when general prescriptions are given to a programer, be it in the development of a film, a tape, or printed materials, a great deal of latitude remains in his hands as to the appearance of the final product in its most minute detail.

VI. OPTIMIZATION EFFORTS

The major investigation of the relationship of student characteristics to performance was conducted during the spring 1970 run of the course. Some background information regarding student characteristics was given on pages 36 to 40 of this report.

Determination of the Analyses

The initial analysis of the relationship of student characteristics to performance involved prediction of final course achievement. This analysis involved the regression of posttest performance on the battery of student variables. This type of analysis provides insight into identifying students unlikely to attain a satisfactory level of achievement. Although this is certainly an important goal in itself, it does not provide direction in how to design and program the instruction in order to optimize performance for each student. Therefore, subsequent analyses involved the investigation of student variables relating to performance with particular media and various presentation forms or conditions of instruction. These analyses were conducted as a subset within the scope of the overall research program investigating group or mean performance as discussed in Chapters I through V.

A total of 44 midshipmen were enrolled in the course. Although a larger number of students might have been desirable, with the statistical controls employed, this number was sufficient for analysis of mean performance for each of the variables investigated. However, certain restrictions were necessary in the analysis of the relationship of student characteristics to performance in the various conditions of instruction. Considering the relatively small number of students, the only regression analyses that could be conducted were those that dealt with the relationship of student characteristics to overall performance on media, and conditions of instruction involving comparisons within subjects, which in both cases would provide data on all 44 students.

As can be seen in Table 3 (p. 58), Experiment I involved sixteen segments in which three variables were manipulated. Only the variable of media (audiotape vs. videotape) was a within student comparison. That is, each student worked through half of the segments with videotape and the other half with audiotape.

Therefore, an analysis of the relationship of student characteristics to performance with audiotape as opposed to videotape could be conducted. In addition, since each of the 44 students used both audiotapes and videotapes across these segments, an analysis of the relationship of student characteristics to performance in taped media (audiotape and videotape combined) was also conducted.

Again referring to Table 3, it can be seen that experiment II involved nine segments in which two variables were manipulated with the medium of linear text being used consistently throughout these segments. Only the variable of the form of the response demanded of the student (overt selected, overt spoken and covert) was a within student comparison. Each of the 44 students worked with each of the three types of response demand. Therefore, in this experiment an analysis of the relationship of student characteristics to performance in each condition of responding as well as to performance with linear text in general was conducted.

Experiment III covered eight segments in which two variables were manipulated. Neither of these variables listed in Table 3 was a within student

comparison. Each student saw only one of the four conditions listed for this experiment, thus leaving only 11 students in each condition. With the large number of student characteristics investigated it was not feasible to conduct regression analysis on this data. However, since all 44 students used an audiotape with an intrinsically programed booklet (AT/IP) in the first four segments and in the other four segments all 44 students worked with computer-assisted instruction (CAI), an analysis of the relationship of student characteristics to performance on AT/IP vs. CAI and branching media in general (a combination of AT/IP and CAI) was possible.

Experiment IV involved nine segments, all using the medium of syndactic text, as indicated in Table 3, in which the type of remediation method was manipulated. This was a within student comparison in which each student studied under each of the three conditions. Therefore an analysis of the relationship of student characteristics to performance in each of these conditions as well as performance with syndactic text was conducted.

Since the variable being investigated in

Experiment V was not a within student comparison and since the decision was made to change the medium used in the three segments involved, no analysis of student characteristics was conducted in this experiment.

In summary, there were 13 basic types of analyses conducted during the spring run relating student characteristics to performance on various media and conditions of instruction as well as to overall performance as measured by the posttest (see Table 6). In all cases but the posttest, the criterion variable or measure of performance used was the cumulative posttest. For each of these conditions of instruction three separate regression analyses were conducted. The student characteristics were analyzed in relation to the acquisition of knowledge (Type I CPT test items), and the application of knowledge (Type II CPT test items) as well as the two types of tasks combined (total CPT items). The classification of these two types of test items roughly corresponds to Bloom's categories of knowledge and applications. (Bloom et al, 1956).

TABLE 6

SUMMARY OF REGRESSION ANALYSES CONDUCTED*

Experiment	Criterion	Segments	Predicted Performance
---	Posttest	1.1-12.4	1)Final Course Achievement
I	CPT-1	2.2-2.5	2)Audiotape vs. Videotape
	CPT-3	3.1-3.4	3)Taped Lecture (Audio & Video combined)
	CPT-7	5.7-5.10	
	CPT-9	7.1-7.4	
II	CPT-4	4.1-4.3	4)Linear Text
	CPT-6	5.4-5.6	5)Overt selected response demand
	CPT-10	8.1-8.3	6)Overt spoken response demand
			7)Covert response demand
III	CPT-5	4.4-4.7	8)Computer-Assisted Instruction (CAI)
	CPT-13	12.1-12.4	vs. Audiotape/Intrinsically Programed Booklet (AT/IPB)
			9)Branching Media (CAI and AT/IP combined)
IV	CPT-2	2.6-2.8	10)Syndactic Text
	CPT-8	6.1-6.3	11)High response demand remediation
	CPT-12	11.1-11.3	12)Low response demand remediation
			13)No remediation

* For each of the conditions of instruction 2 through 13, three separate regression analyses were conducted. The student characteristics were analyzed in relation to the acquisition of knowledge, the application of knowledge, and the two types of tasks combined as measured on the Cumulative Posttest (CPT).

As can be seen in Table 6, analyses 2, 3, 4, 8, 9 and 10 involved an investigation of the relationship of student characteristics to performance within a particular medium. Analyses 5, 6 and 7 involved the relationship of student characteristic to performance within linear text, but specifically to the conditions of instruction where the response required of the student was varied. In analyses 11, 12 and 13, the relationship of student characteristics to performance on a particular form of remediation (or lack of it) within syndactic text presentations was investigated.

A syndactic text is essentially a series of linear programed frames each preceded by a brief but complete summary of the information presented in the frames. Students worked through the syndactic text by reading the first summary statement and taking a summary quiz of five to eight questions. If the student answered all summary quiz questions correctly, he read the second summary, took summary quiz 2, etc. The student who incorrectly answered one or more questions of a summary quiz was required to remediate through the linear programed sequence associated with that summary.

Method

Test Battery. A battery of 137 predictor variables was used in the regression analyses. Included in the battery were common standardized tests in the major areas of aptitude, achievement, personality, motivation, and interest. Also included were items of student questionnaire data. Emphasis in the selection of tests was on commonly used and well-standardized tests, with considerable established validity to aid in the interpretation of findings. Emphasis in the student questionnaire items was on face validity.

In addition to such achievement variables as cumulative grade point average, converted rank in class, and high school recommendation score, the battery included the SAT-Verbal, SAT-Math, CEEB English Comprehension, CEEB Math Achievement and the various scales of the Edwards Personal Preference Schedule, the 16 Personality Factor Scale, the Ohio State Psychological Examination, the Strong Vocational Interest Blank and the 22 questions on the Student Questionnaire. The Student Questionnaire dealt with topics such as high school or college subjects studied, methods of previous instruction, study habits and college related abilities. A complete listing of the predictor variables is given in Appendix A and the complete Student Questionnaire is given in Appendix B.

Because of the large number of predictors and the small number of students available, and the fact that little confidence could be placed in most a priori hypotheses relating performance and predictors, the analyses of Phase II of the USNA Leadership course development project were designed as a variable selection process. The aim was to filter out potential important predictors from the many candidates available.

Criterion Variables. Three types of dependent measures were used as the basis of the multiple regression analyses. First was the administrative posttest used as the criterion variable for prediction of overall course achievement. The second type of criterion variable was the student total residual derived from average student performance in each condition of instruction, which was used as the criterion variable in prediction of achievement with a particular medium. The third type of criterion variable was the within-student residual derived from scores on a module, used as the criterion variable in predicting achievement in a particular presentation form or condition of instruction. The latter two types of criterion variables are identified as sources

of error variance in the analyses of variance and represent unexplained individual differences in student performance after overall treatment conditions and Cumulative Posttest (CPT) unit differences are removed. In every experiment, residuals were derived for total CPT scores, CPT Type I scores, and CPT Type II scores.

A total residual was obtained from a student's mean performance over all CPT units of an experiment by subtracting the mean of the group (to which the student belongs in that experiment) from the student's mean. The resulting deviation score represents how well the student learned in relation to his group over the entire experiment. Since each experiment involved a particular medium, this score indicates how well the student learns in connection with that medium, at least for the kinds of content and presentations used with the medium. Regression of the tota. residuals on the battery of student variables could thus be used to identify variables associated with variation in achievement with particular media.

A within-student residual was derived by subtracting the mean for the student's group in a particular condition of instruction from the student's score in that condition, and secondly, subtracting the total residual for the student from the result

of the first subtraction. The resulting deviation score represents how well the student learned in relation to his average standing in the group, and in relation to the average performance of the group in that particular condition. When the within-student residuals for a particular condition of instruction are regressed on the battery of student variables, variables are identified predicting performance in the presentation conditions defining that condition.

Preliminary Variable Selection

The analyses for each criterion variable were conducted in three stages. The first stage involved the identification of potential predictor variables for input to the step-up multiple linear regression analysis. The following rules were employed in selecting these variables from the total pool of 137 student variables. A variable was selected if its first-order correlation with the criterion was .20 or greater. For each of the primary variables selected according to this first rule, its major correlate was included in the step-up regression analysis if it correlated less than .20 with the criterion variable but .40 or greater with the primary predictor. This latter rule was intended to select possible suppressor variables. In addition, 15 preselected predictor variables

were added if they were not included according to the above rules. The 15 preselected predictor variables were those that have commonly been used in predicting course achievement, and were preselected variables in the regression analyses in order to give them maximum opportunity to demonstrate their predictive power. These 15 variables are identified in Appendix A.

Step-Up Regression Analysis. The second stage of each analysis involved the input of the potential predictor variables identified in the preliminary variable selection process to a step-up regression analysis. The step-up multiple regression analysis involves the computation of a sequence of multiple linear regression equations in a stepwise manner. At each step one variable is added to the regression equation. The variable added is the one which makes the greatest reduction in the error sum of squares. Equivalently, it is the variable which has the highest partial correlation with the dependent variable partialed on the variables which have already been added. This amounts to being the variable which, if it were added, would have the highest F value. The computation was set to stop when the F value for a variable was not significant at the .10 level or less.

Step-Down Regression Analysis. In the final stage of the analyses, the variables surviving the step-up regression analysis served as input to the step-down regression analysis. In essence the step-down analysis is a reversal of the step-up analysis. It involves the computation of a sequence of regression equations in a stepwise manner. At each step it selects the variable with the smallest computed t value and looks at it as though it were the last variable entered. If this variable does not make a significant reduction in the error sum of squares, it is dropped from the analysis and the t values for the remaining variables are recomputed and the process is repeated. The accepted significance level was set at .01. When a predictor variable is significant at this level (when the loss in prediction dropping that variable is significant at the .01 level), the computation stops. All the remaining variables are significant predictors of performance in the particular condition being investigated. Procedures of the step-up and step-down analyses are based on those described in Draper and Smith (1966).

Discussion of the Spring 1970 Analyses

In analyzing the relationships that have been found it was recognized that there may be several different interpretations of why a particular variable relates to a particular medium or particular condition of instruction. Therefore, rather than going into an indepth discussion for each variable in every analysis, the discussion will concern itself with identifying general classes or clusters of variables that appear to relate to performance within the analyses and, where possible, to identify differences across the various analyses conducted.

In interpreting the reported relationships of student characteristics to the various media and conditions of instruction the following factors should be kept clearly in mind: 1) the instructional system, 2) the content being taught, 3) the medium used, and 4) the variations of the conditions of instruction within and across each medium. The instructional system basically required the student to proceed with a segment of instruction programed in a particular manner, and then to take a criterion referenced progress check. If he achieved 80% or better, he could proceed to the next

segment of instruction. If he failed to achieve that level of performance, he was required to remediate the identified areas of deficiency using specific review materials. The research embedded in the course required the student in most cases to delay remediation over several segments involving a particular research question. Also the research involved in no way hindered any student's final performance. It did require him to follow specific procedures that varied from one unit of instruction to another. With respect to the content, a perusal of Table 2 indicates that many of the topics covered in the course are interrelated, but that there is a diversity of content area taught. A wide variety of media was utilized across this diversity of content area. In addition, specific conditions of instruction were employed within these media. These factors were taken into account in designing the research involving group comparisons of performance for specific conditions in instruction within media.

It should be noted that some problems of interpretation of the analyses of student characteristics can arise if these factors are not kept in mind. In interpreting the relationship of the various aptitude, personality, and self-interest and self-report variables, it may be difficult to determine the relative effect of the system, the content, the media, and the conditions of instruction. The point is that some of these relationships may be obvious while many may not be.

The prediction of final course achievement will be discussed first followed by the relationships of student characteristics to performance with the various media utilized. The section will conclude with a discussion of the predictors of performance and the various conditions of instruction.

Final course achievement. The prediction of final course achievement, it should be recalled, was different from the remainder of the analyses in that the criterion variable was performance on an 80-point criterion-referenced test and it did not involve separate analyses for different types of tasks. The interpretation of this

analysis is rather straightforward. The variables predicting posttest performance were quite diverse. In addition to the variables of prior knowledge (pretest) and general ability (English comprehension), there were three personality variables (achievement, autonomy, and humble vs. assertive), one occupational self-interest variable (pharmacist), and one self-report variable (average hours of study).

In an individualized course stressing a pre-set level of performance for each student, it would be somewhat surprising to find the pretest as a predictor of final course achievement if the instructional materials and tests had been completely validated. It should be noted that this data was tabulated on the basis of the first full scale implementation of the materials. Finding the pretest as a predictor does indicate a need for revision of materials and tests. In fact this revision cycle was planned.

It would appear that individuals who score high on the final examination tend to have good reading aptitude, particularly comprehension, which may be related to test taking ability. This may well account for the negative relationship of number of hours typically

spent studying. Personality characteristics found as predictors indicate that these individuals tend to be assertive, self-assured and independent minded yet do not avoid responsibilities and obligations, or rebel against authority. The negative relationship of achievement as measured on the Edwards Personal Preference Schedule indicates that these students are not highly motivated to accomplish tasks requiring skill and effort, or to do a difficult job well. It may well be, however, that they simply did not perceive the course as difficult or something that required great skill and effort. The relationship of interest in the profession of pharmacist as measured on the Strong Vocational Interest Blank may indicate an interest in attending to small details which the profession of pharmacist certainly requires. There are indeed many details to be attended to in the individualized multi-media leadership course if a high level of performance is to be achieved.

Media predictors. When looking at the predictors resulting from the step-down analyses for the different types of tasks measured on the CPT's within media as well as across media, no clear pattern emerged. However, when the step-down analyses were supplemented with the step-up analyses and the first-order correlations,

the predictors of overall performance under different media seemed generally consistent. There appeared to be a cluster of verbal skill variables such as CEEB English comprehension, reading comprehension and total reading from the Ohio State Examination, and the SAT-Verbal that were always related to overall performance regardless of the media or type of task involved. In addition to the standard variables which one might expect to find, there appeared to be a cluster of variables related to performance that was unique to each of the media involved.

In the case of audiotape and videotape this second order cluster of variables was also in the verbal skills area, but it was more related to oral expression rather than reading and test taking ability. Several of the self-rating student report variables from the student questionnaire appeared to relate to performance with the taped media. The self-report variables of previous instruction by audiotape, and college-related abilities with respect to vocabulary, reading, writing, and oral expression all show up in the first order correlation. These variables did not appear in linear text or syndactic text as they did with the taped lecture media. These self-report variables relate to ability to learn from oral presentations which, of course, is involved in both audiotape and videotape.

In contrast to the auditory learning cluster related to the taped lecture media, more frequent correlations with different personality and self-interest variables appeared in relation to linear text and syndactic text. With respect to linear text, the first order correlations showed negative relationships for shy vs. outgoing and exhibitionism, while interest in the profession of librarian was positively related to performance. It was generally the case that interest in psychology, musician performer, and music teacher for example, which are more related to public exhibitionism of products of work, were negatively related to the shy vs. outgoing, exhibitionism and librarian types of scales. Therefore, it would appear that there was a general introversion-extroversion cluster of variables that was involved in performance with linear text, where the more withdrawn type of personality achieved a higher level of performance. Some of the other interest and self-report variables may be as much related to the particular content as to the medium in which it was programed. The strong relationship of converted rank in class with linear text may have been more related to motivation to study than to academic skills.

With respect to syndactic texts, another type of personality dimension relating to performance was found. This was the variable of conservative vs. experimental, where a higher level of performance was achieved for the experimental personality. This type of individual is more inclined to experiment in life generally and is more tolerant of inconvenience and change. It would appear that the novelty of syndactic texts was more readily adaptable to individuals with an experimental personality trait. As with linear text, there were a variety of self-interest variables that may be related to the content as well as to the medium itself.

The secondary cluster of variables relating to performance with computer-assisted instruction (CAI) and its parallel, audiotape with an intrinsically programed booklet (AT/IPB), was perhaps the most difficult to clearly identify. Although there are some consistencies, there were, in these analyses, a variety of personality variables and self-interest variables that were difficult to reconcile when going from Type I tasks to total CPT performance and when looking at the analysis of the CAI-AT/IPB differences as compared to the two media combined. It should be noted that these media were actually composite media. In addition, the experimental conditions

were slightly different for these media than for the others. In all other cases, the experimental conditions involved within student comparisons where all students saw each of the varied conditions within the medium. The analysis of the experimental conditions within CAI and AT/IPB were between subject comparisons. These factors may be contributory to the lack of clear findings.

Conditions of instruction predictors. The relationships of student characteristics to the conditions of instruction involving variations in the response demanded of the student and the type of remediation appeared to be different from the analyses involving overall performance on media. This was the case even though the response demanded was varied within the medium of linear text and the remediation type of variable within the medium of syndactic text. In general, the verbal skills cluster of variables did not appear. In the main there were a variety of personality, self-report and self-interest variables that appeared with no consistent pattern except perhaps for the overt selected and covert response demand forms. However, in this case, the finding that a concrete thinker would perform better with the

covert response demand and an abstract thinker would do better with the overt selected form seems somewhat intuitive. There were some procedural problems in implementing these conditions that causes the reliability of these particular findings to be questioned. The students generally reported that they did not always strictly follow the instructions. With respect to the remediation type, the students performed so well on the syndactic text summaries that many did not need the remediation at all.

Conclusions Relating to the Spring 1970 Data

In addition to finding significant predictors of final course achievement relating to aptitude, personality, and interest, this investigation identified the general cluster of verbal skill variables that related to performance regardless of the media involved. The fact that in general there were no particular variables or group of variables that were uniquely related to performance on a lower level learning task as opposed to a higher level learning task may be a reflection of too broad a classification of types of learning tasks. In addition to the cluster of verbal skill variables that relate to performance regardless of the media employed, a secondary cluster of variables was found that was generally unique to performance within each of the media involved.

Although there were some procedural problems in the implementation of the course that caused some difficulty in interpreting some of the analyses of the student characteristics, the methodology appears rather sound, and the identification of general clusters of variables was of definite value.

While it is not recommended that these findings be applied in an ongoing course until they are cross validated, the more reliable findings could be used to tentatively identify individuals who might have problems learning from a particular medium. In the Leadership, Psychology and Management course this would entail the determination of an acceptable base level of performance on the norm-referenced cumulative posttest and the determination of the relationship of these tests to overall performance in the course. It is felt, however, that the maximum benefit to be gained from this effort, particularly without cross validation, is in providing insight and direction for future research and application of the relationship of student characteristics to performance in individualized multi-media course presentations.

Student Characteristics - the Fall run

The investigation of student characteristics in the fall 1970 implementation of the course was conducted on a much smaller scale than was done in the spring run. The basic strategy was to select variables from each of the analyses conducted with the spring 1970 data that would appear to have the highest payoff in the fall run and that would have some applicability for the permanent course. Ten major variables were selected and correlations of each of these variables were conducted with performance on all of the cumulative posttests (CPT's) except one. The one exception was the CPT data covering segments developed with the medium of Computer Assisted Instruction.

Table 7 presents the results of these correlations. The independent measures were student's scores on:

(1) The SAT-Verbal; three scales from the Edwards Personal Preference Schedule

(2) Achievement

(3) Exhibition

(4) Autonomy; five scales from the 16 Personality Factor Questionnaire

Table 7

Fall Run Student Characteristics Correlations

Dependent Measures	Independent Measures									
CPT.	1	2	3	4	5	6	7	8	9	10
1	.44 (16)	.36 (14)	.21 (14)	.11 (14)	.01 (14)	.24 (14)	.09 (14)	.02 (14)	.18 (14)	.47 (14)
3	XX.62 (19)	.09 (17)	.20 (17)	.41 (17)	.04 (16)	.10 (16)	.41 (16)	.10 (16)	.40 (16)	.18 (13)
7	.10 (19)	.18 (17)	.03 (17)	.45 (17)	-.28 (16)	-.30 (16)	-.13 (16)	.13 (16)	-.00 (16)	.47 (13)
9	.32 (16)	.13 (14)	-.20 (14)	.36 (14)	-.18 (14)	-.12 (14)	.18 (14)	-.15 (14)	-.35 (14)	.23 (14)
4	-.01 (18)	-.30 (16)	.41 (16)	.45 (16)	-.32 (15)	.28 (15)	X.63 (15)	.16 (15)	.49 (15)	.22 (14)
6	.44 (18)	.45 (16)	.02 (16)	.22 (16)	.01 (15)	-.27 (15)	-.02 (15)	.02 (15)	.20 (15)	.40 (14)
10	-.12 (18)	-.02 (16)	.15 (16)	.39 (16)	.08 (15)	X.52 (15)	.05 (15)	-.30 (15)	XX.66 (15)	.49 (14)
5A	-.21 (8)	-.26 (6)	.47 (6)	.28 (6)	-.63 (5)	.09 (4)	-.35 (5)	-.23 (5)	.21 (5)	.77 (5)
5B	.64 (7)	.68 (5)	X-.89 (5)	-.60 (5)	.41 (5)	.60 (6)	.49 (6)	.38 (6)	.86 (6)	.09 (5)
2	.11 (24)	.16 (21)	.03 (21)	.18 (21)	-.05 (20)	-.11 (20)	-.05 (20)	.31 (20)	.33 (20)	.15 (17)
8	XX.64 (22)	.38 (20)	-.15 (20)	-.19 (20)	.16 (20)	.23 (20)	.09 (20)	.30 (20)	-.05 (20)	X.59 (17)
12	.45 (24)	.11 (21)	-.34 (21)	.22 (21)	.13 (21)	-.00 (17)	.05 (20)	.04 (20)	.11 (20)	.33 (19)
14	.23 (34)	.28 (30)	.05 (30)	-.08 (29)	.04 (29)	-.09 (29)	.00 (29)	-.13 (29)	-.06 (29)	.18 (26)
11	.24 (34)	-.07 (29)	-.13 (30)	-.05 (29)	-.07 (29)	.04 (29)	-.24 (29)	.13 (29)	.28 (29)	.25 (26)

* significant at the .05 level

** significant at the .01 level

The number of students involved in each analysis is given in parentheses after the correlation.

(5) Reserved vs. outgoing

(6) Concrete thinking vs. abstract thinking

(7) Humble vs. assertive

(8) Conservative vs. experimenting

(9) Leadership

(10) The Ohio State Psychological reading comprehension section.

Each of these variables was correlated with performance on fourteen separate measures. The first four measures were performance on:

(1) CPT 1

(2) CPT 3

(3) CPT 7

(4) CPT 9

Each of these measures reflected performance with the medium of audiotape/panel book. Only those students who went through the high RDF conditions were included in this analysis. This restriction was set as it was for each of the other measures so that the results would have maximum applicability with respect to the final configuration of the course as it was implemented in the installation run. Thus, the number of students involved in these analyses ranged from 13 to 19. The number of

students in each case is given in parentheses next to the correlation. The next three measures [(5) CPT 4, (6) CPT 6, and (7) CPT 10] dealt with performance with linear text. Again only those students who went through the high RDF conditions were selected. The next two measures reflected performance with the medium of (8) audiotape with an intrinsic booklet and (9) the script of the audiotape with the intrinsic booklet. For CPT 5(A) only the data for the students who went through the high (RDF)--Medium (MF) were used, and for CPT 5(B) only students who used the script version. The next three measures [(10) CPT 2, (11) CPT 8 and (12) CPT 12] dealt with performance with the syndactic text medium where either the linear sequence or the detailed summary was available as internal remediation. The last two measures [(13) CPT 14 and (14) CPT 11] reflected performance with the medium of syndactic text where the only internal remediation was the linear sequence.

As can be seen from Table 7 there were few significant results. The SAT-Verbal showed a significant and positive correlation with one of the audiotape/panel book units and one of the syndactic text units. An additional variable, the reading comprehension section of the Ohio State Psychological test also showed a significant

positive correlation with the same unit of syndatic text. This was one of only two cases where more than one of the variables showed a significant correlation within the same unit of content. These two tests are basically measures of reading comprehension and it is quite reasonable that they would correlate with performance using syndactic text. In this medium the material is presented in printed form in larger chunks than with any other medium used. With respect to performance with linear text, three personality variables showed significant positive correlations with two of the content units in which linear text was employed. Each of these variables is from the 16 Personality Factor Questionnaire. A positive correlation with the first one of the three, concrete thinking vs. abstract thinking, indicates that a person who tends to be quick to grasp ideas, a faster learner, tends to perform well with linear text. The second variable (humble vs. assertive) indicates that the individual who performs well with linear text is also assertive, self-assured and independent minded. In addition, the positive correlation with the leadership scale indicates this individual has characteristics similar to those of effective leaders.

The only other significant correlation was the negative relationship of exhibition with performance with

a tape script and intrinsic booklet. In essence this means that the individual who performs well with tape script and intrinsically programed booklet does not care to be the center of attention.

VII. SUMMARY OF EVALUATION RESULTS

The Installation Run

The initial implementations of the course in the spring and fall semesters of 1970 generally validated subsystems and the basic plan of operation for the Leadership, Psychology and Management course. In essence, the final course system which was installed in the spring of 1970 reflected successive refinements made on the basis of the two previous administrations.

In arriving at the final course configuration a variety of inputs were considered. The empirical data considered were the research and effectiveness data, student ratings and time. Two equally important inputs were administrative ease of use and costs. In lieu of significant and conclusive research findings the later two inputs were given considerable weight. In general, the high response demand frequency version with overt selected responses for each segment was selected. A syndactic text or script version for each of the hardware bound media was made available as an alternate version for the installation run in the spring 1971 semester.

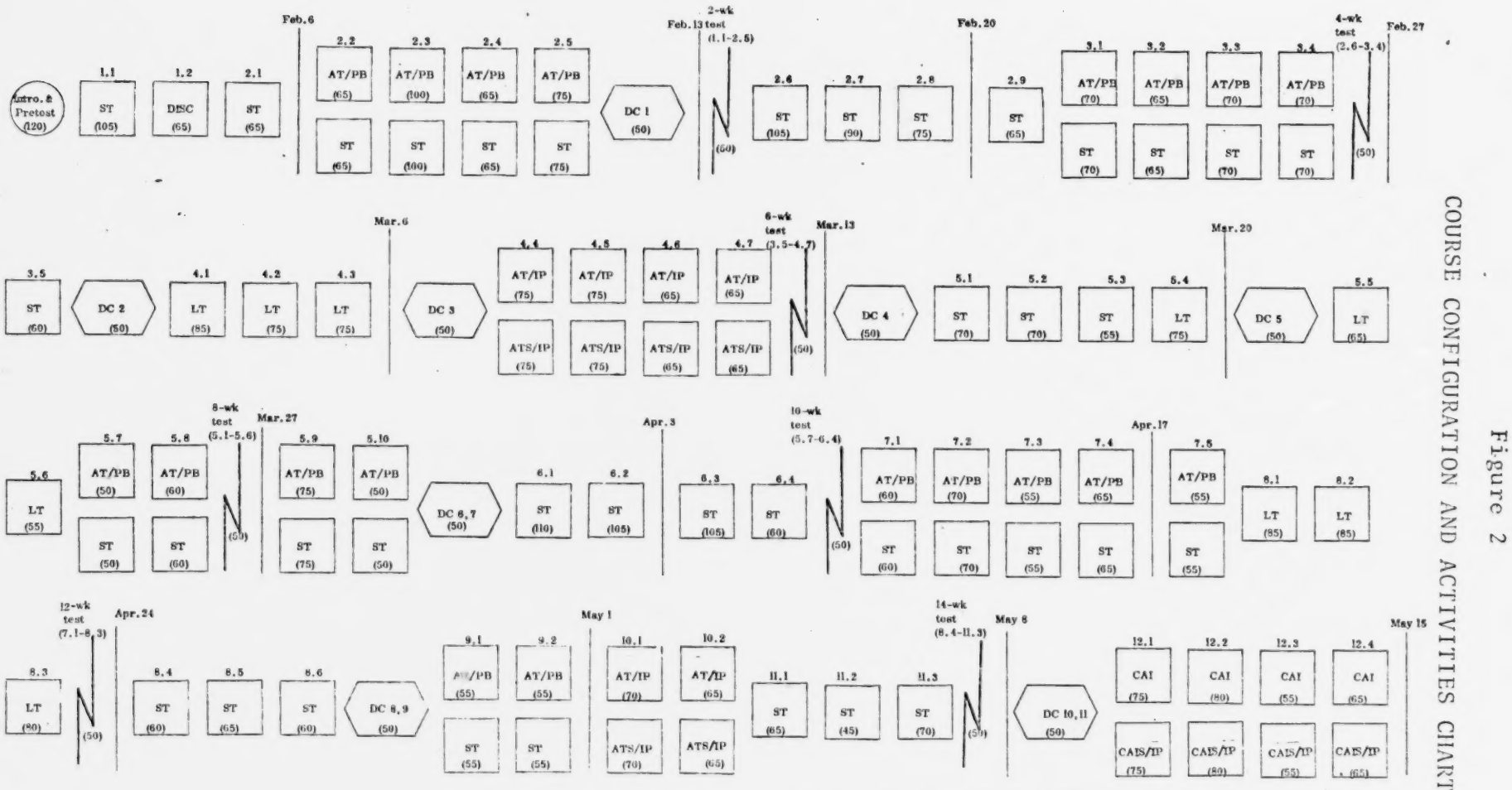
Figure 2 presents the course configuration and activities chart. It can be seen that 29 of the 59 segments involve a media option. Of the remaining 30 segments, 20 are programed in the medium of syndactic text which provides the student with an option in terms of management. Table 8 indicates the packaging of the segments by volume and the alternate media available.

Figure 2

COURSE CONFIGURATION AND ACTIVITIES CHART

INTRODUCTION TO PSYCHOLOGY AND LEADERSHIP

SPRING 1971



LEGEND

AT/IP = Audiotape/Intrinsically Programed Booklet
ATS/IP = Audiotape Script/Intrinsically Programed Booklet
AT/PB = Audiotape/Panel Book
CAI = Computer Assisted Instruction
CAIS/IP = CAI Script/Intrinsically Programed Booklet
DC = Depth Core Sessions
DISC. = Group Discussion
LT = Linear Text
ST = Syndactic Text
( ) = Estimated time of Activity

Admin. Test

Completion dates of minimum units of materials

TABLE 8

SEGMENTS BY VOLUME AND MEDIA EMPLOYED

Volume Number	Segments Included	Prime Media	Alternate Media	Volume Number
I	1.1	Syndactic text		
	1.2	Discussion booklet		
II-A	2.1	Syndactic text	Syndactic text	II-B
	2.2	Audiotape/Panel book	Syndactic text	
	2.3	Audiotape/Panel book	Syndactic text	
	2.4	Audiotape/Panel book	Syndactic text	
	2.5	Audiotape/Panel book	Syndactic text	
II-C	2.6	Syndactic text		
	2.7	Syndactic text		
	2.8	Syndactic text		
	2.9	Syndactic text		
III-A	3.1	Audiotape/Panel book	Syndactic text	III-B
	3.2	Audiotape/Panel book	Syndactic text	
	3.3	Audiotape/Panel book	Syndactic text	
	3.4	Audiotape/Panel book	Syndactic text	
	3.5	Syndactic text	Syndactic text	
IV-A	4.1	Linear text		
	4.2	Linear text		
	4.3	Linear text		
	4.4	Audiotape/Intrinsically Programed booklet	Tape Script & Intrinsically Programed booklet	IV-Script
IV-B	4.5	Audiotape/Intrinsically Programed booklet	Tape Script & Intrinsically Programed booklet	
	4.6	Audiotape/Intrinsically Programed booklet	Tape Script & Intrinsically Programed booklet	
	4.7	Audiotape/Intrinsically Programed booklet	Tape Script & Intrinsically Programed booklet	
V-A	5.1	Syndactic text		
	5.2	Syndactic text		
	5.3	Syndactic text		
	5.4	Linear text		
	5.5	Linear text		
	5.6	Linear text		
V-B	5.7	Audiotape/Panel book	Syndactic text	V-C
	5.8	Audiotape/Panel book	Syndactic text	
	5.9	Audiotape/Panel book	Syndactic text	
	5.10	Audiotape/Panel book	Syndactic text	

TABLE 8 (Continued)

Volume Number	Segments Included	Prime Media	Alternate Media	Volume Number
VI-A	6.1	Linear text		
	6.2	Linear text		
VI-B	6.3	Linear text		
	6.4	Syndactic text		
VII-A	7.1	Audiotape/Panel book	Syndactic text	VII-B
	7.2	Audiotape/Panel book	Syndactic text	
	7.3	Audiotape/Panel book	Syndactic text	VII-C
	7.4	Audiotape/Panel book	Syndactic text	
	7.5	Audiotape/Panel book	Syndactic text	
VIII-A	8.1	Linear text		
	8.2	Linear text		
	8.3	Linear text		
VIII-B	8.4	Syndactic text		
	8.5	Syndactic text		
	8.6	Syndactic text		
IX-A	9.1	Audiotape/Panel book	Syndactic text	IX-B
	9.2	Audiotape/Panel book	Syndactic text	
X	10.1	Audiotape/Intrinsically Programed booklet	Tape Script & Intrinsically Programed booklet	X Script
	10.2	Audiotape/Intrinsically Programed booklet	Tape script & Intrinsically Programed booklet	
XI	11.1	Syndactic text		
	11.2	Syndactic text		
	11.3	Syndactic text		
	12.1	Computer-Assisted Instruction	Intrinsically Programed booklet	XII-A
	12.2	Computer-Assisted Instruction	Intrinsically Programed booklet	XII-B
	12.3	Computer-Assisted Instruction	Intrinsically Programed booklet	XII-C
	12.4	Computer-Assisted Instruction	Intrinsically Programed booklet	XII-D

Just as the course content and structure underwent successive refinements based on the previous administrations of the course, so too did the manner in which the student interacted with the instructional materials. The refinements were made on the basis of the requirements for revision and research. Basically, each student studied a segment of instruction and then was tested on his mastery of the objectives for that segment. If he achieved 80 percent or better on this criterion referenced progress check, he proceeded to the next segment. If he did not achieve this level, he was given specific remediation instructions and was required to retake the progress check. If he still failed to achieve the appropriate level of mastery, he received special tutoring from the instructor. To obtain maximally reliable data for revision from the first implementation, the midshipmen were required to take the progress checks individually in class under the guidance of the course administrator. In essence, this required the midshipmen to come in after each segment. To allow the midshipmen more freedom in the fall run, they were given the progress checks with the materials but they still had to come in and have the course administrator grade them to determine if they needed remediation.

The research imbedded in the course imposed further restrictions on the students' interaction with the instructional materials. In order to accurately assess the research questions being asked, separate norm-referenced

tests (Cumulative Posttests) were developed which the students had to take in class at specific points in the instruction. The students also had to delay remediation on segments within a Cumulative Posttest Unit until they completed the research test. However, with the completion of the research in the fall run, the Cumulative Posttests and the resultant restrictions were removed from the course.

The installation run of the course in the spring of '71 was characterized by maximum freedom for each student to proceed at his own pace and to manage his own time and place of study. The freedom for the student to manage his own instruction is an important motivational effect of an individualized instructional system. The rather rigid controls imposed on the students in the previous implementations of the course, which were necessary in order to obtain valid revision and research data, were no longer necessary. With the removal of the research Cumulative Posttests, the students were required to come to scheduled classes only for depth core sessions and administrative tests (see Figure 2). The progress checks, which were controlled on site, were used as self tests for the students.

Thus, the students graded their own progress checks and determined their requirements for remediation and/or tutoring with the instructor. The instructor was, of course, available for any non-required tutoring and for assignment

of enrichment sessions. Rather than being required to prepare thrice-weekly lectures, the instructor was able to concentrate on the integration and transfer of important concepts, and working with individual students as they needed his assistance. In addition, the instructor was aided by the course administrator in record keeping and scheduling.

Comparative Analyses of the Effectiveness of the Three Implementations

In judging adequacy of segment performance, WLC established a criterion of 80 - 80, that is, 80% of the students must correctly answer 80% of the progress check test items, if not on the first attempt (before remediation), then on the second attempt (after remediation) or in individual tutoring sessions. Table 9 reports the percentage of students achieving the 80% criterion level before remediation on each of the three course implementations. The mean percentages in this table show a vast improvement in overall segment performance from 80% - 57% in the first run, to 80% - 82% in the second run, and 80% - 94% in the final run. These data reveal that the requirement of the two revision cycles to reduce the need for remediation and tutoring was amply satisfied. In the third run there are only four segments (2.3, 2.6, 2.7, 3.4) which do not meet the 80 - 80 criterion before remediation. It should be pointed out that in three of these segments the number of test items was 8 or 9 rather than the customary 10. Actually only one segment (Segment 2.3) in the final run fell below the 80 - 80 criterion based on a 10-item progress check, and this was a marginal drop (77%). In the final run all students achieved at least 80% correct or better after remediation, and there were no instances in which tutoring was necessitated. Thus, in assessing the final run of the

Table 9

Percentage of Students, by Segment, Attaining 80% or Better on Progress Checks Before Remediation, Across Three Implementations

Segment Number	First Run	Second Run	Third Run
1.1	70.5	91.1	94.9
1.2	75.0	80.0	85.7
2.1	65.9	86.7	95.9
2.2	31.8	85.7	91.8
2.3	18.2	35.7	76.6
2.4	34.9	69.0	88.8
2.5	25.0	88.1	97.9
2.6	40.9	68.3	69.4
2.7	11.4	42.9	73.5
2.8	86.4	88.1	93.9
2.9	36.4	90.5	96.9
3.1	22.7	40.5	80.4
3.2	20.5	69.0	86.6
3.3	63.6	40.5	100.0
3.4	72.1	88.1	73.5
3.5	83.7	95.2	90.8
4.1	61.4	95.2	97.9
4.2	88.6	100.0	96.8
4.3	88.6	83.3	94.8
4.4	95.5	97.7	98.9
4.5	43.2	100.0	96.9
4.6	52.3	97.7	97.9
4.7	68.2	95.5	95.9
5.1	43.2	90.9	90.8
5.2	22.7	40.9	86.7
5.3	52.3	95.5	97.9
5.4	25.0	75.0	93.9
5.5	4.5	72.7	96.9
5.6	13.6	79.5	87.8
5.7	52.3	81.8	96.9
5.8	77.3	86.4	93.8
5.9	39.5	56.8	93.8
5.10	20.9	95.5	91.8

Table 9 (Continued)

Segment Number	First Run	Second Run	Third Run
6.1	50.0	95.5	95.9
6.2	72.7	81.8	96.9
6.3	95.5	86.4	94.9
6.4	79.5	95.3	95.9
7.1	18.2	69.8	92.6
7.2	45.5	90.7	97.9
7.3	34.1	69.8	92.6
7.4	59.1	60.5	93.7
7.5	43.2	76.7	98.9
8.1	70.5	86.0	93.7
8.2	42.9	97.7	97.9
8.3	88.6	97.7	100.0
8.4	43.2	97.7	100.0
8.5	60.5	88.4	92.4
8.6	56.8	79.1	95.7
9.1	68.2	93.0	98.9
9.2	93.2	93.0	98.9
10.1	90.9	97.7	98.9
10.2	95.5	97.7	100.0
11.1	70.5	74.4	97.8
11.2	88.6	100.0	100.0
11.3	52.3	95.3	81.5
12.1	84.1	97.7	100.0
12.2	59.1	74.4	98.9
12.3	97.7	93.0	96.7
12.4	84.1	67.4	98.9
Mean, All Segments	56.8	82.4	93.5

course by percent-percent standards, 55 segments meet or greatly exceed the 80 - 80 criterion before remediation, and only 4 segments were below this level. After remediation, all segments met the 80 - 80 criterion.

Table 10 lists by segment the mean progress check performance before remediation for the three runs of the course. Once again the reader will note the improvement in course performance effected by each revision cycle, such that the mean performance increased by 8 and 6% respectively with each run, from 78% to 86% to 92%. In addition, it can be noted that in terms of overall mean performance there were only ten segments that showed a mean below 90%.

Time data across the three runs are reported in Table 11. The figures in the first and second runs represent the number of minutes spent on instructional materials as well as taking the progress check and performing remediation as needed. The time data for the third run are thus inflated by approximately ten minutes additional to take the progress check and remediate. When the mean time figure from the third run is made comparable to the time expenditures from the first two runs (by subtracting ten minutes from the average of 49 minutes), it is apparent that the average amount of time per segment that the student invests in the course is lowest on the third run, while performance is highest on the third run, indeed, a very desirable combination.

TABLE 10

Segment Progress Check Performance Across Three Implementations
Mean Percentage Correct, Before Remediation

Segment Number	First Run	Second Run	Third Run	Segment Number	First Run	Second Run	Third Run
1.1	82	87	91	6.1	80	90	91
1.2	67	84	87	6.2	82	85	93
2.1	78	88	91	6.3	91	88	92
2.2	68	86	90	6.4	87	87	95
2.3	65	70	85	7.1	69	81	91
2.4	71	78	87	7.2	83	87	97
2.5	65	90	94	7.3	77	80	91
2.6	71	82	86	7.4	78	78	91
2.7	63	72	81	7.5	74	83	95
2.8	89	86	91	8.1	82	87	92
2.9	65	90	91	8.2	81	91	93
3.1	66	73	84	8.3	92	92	96
3.2	66	80	88	8.4	76	94	95
3.3	75	70	95	8.5	77	87	93
3.4	79	87	88	8.6	77	87	94
3.5	84	93	93	9.1	80	91	94
4.1	83	92	96	9.2	88	91	95
4.2	85	95	97	10.1	93	95	95
4.3	87	84	92	10.2	89	91	95
4.4	92	95	96	11.1	80	82	93
4.5	82	95	94	11.2	91	93	97
4.6	73	90	93	11.3	83	90	92
4.7	77	90	93	12.1	87	93	97
5.1	78	89	88	12.2	85	82	94
5.2	70	76	85	12.3	94	93	96
5.3	74	92	93	12.4	85	83	94
5.4	69	84	90				
5.5	59	82	94				
5.6	70	82	92	MEAN-			
5.7	80	83	95	ALL			
5.8	90	90	93	SEGMENTS	78	86	92
5.9	78	79	90				
5.10	66	90	94				

Table 11

Mean Time by Segment Across Three Implementations

Segment Number	Mean Time (No Minutes)		
	First Run	Second Run	Third Run
1.1	92	90	106
1.2	30	48	56
2.1	37	51	53
2.2	31	48	61
2.3	31	76	57
2.4	25	49	51
2.5	26	58	52
2.6	35	68	68
2.7	30	60	67
2.8	34	60	62
2.9	55	52	48
3.1	37	50	42
3.2	25	43	42
3.3	32	48	41
3.4	27	50	44
3.5	34	47	44
4.1	39	61	65
4.2	43	52	41
4.3	41	52	60
4.4	56	67	63
4.5	50	59	53
4.6	41	50	49
4.7	43	48	41
5.1	57	55	52
5.2	59	57	51
5.3	39	41	48
5.4	51	56	51
5.5	52	46	50
5.6	40	40	35
5.7	31	29	38
5.8	39	39	37
5.9	55	54	44
5.10	35	34	32

Table 11 (Continued)

Segment Number	Mean Time (No Minutes)		
	First Run	Second Run	Third Run
6.1	48	67	57
6.2	47	63	53
6.3	45	60	49
6.4	41	44	29
7.1	36	41	41
7.2	42	46	44
7.3	38	38	28
7.4	42	45	41
7.5	39	39	34
8.1	59	64	72
8.2	60	66	63
8.3	52	57	60
8.4	63	44	41
8.5	55	50	47
8.6	56	50	40
9.1	42	42	35
9.2	40	42	37
10.1	52	61	56
10.2	45	50	48
11.1	31	42	38
11.2	30	26	31
11.3	34	46	43
12.1	68	53	60
12.2	71	61	64
12.3	38	36	48
12.4	41	42	54
MEAN	44	51	49

NOTE: Data for the first and second run exclude time spent on taking the progress check and performing remediation. This extra time expenditure of approximately 10 minutes is included in the time data for the third run.

It should be pointed out here that differences in segment performance can be attributed to a multitude of factors such as differences in content, test items, media and presentation forms and the personnel who developed the materials. Thus when differences in segment performance arise, the difficulty lies in determining the causative factors. The one factor which may most easily be isolated is that of the medium employed. To obtain some estimate of the influence the medium itself has on segment performance, one may average the results for material developed in each medium and contrast the averages. This has been done in Table 12.

One important qualification must be made in interpreting the results in Table 12. Although the materials have been grouped on the basis of media, the results should not be construed as evidence of the superiority or inferiority of one medium vis a vis another. These results do not reflect inherent qualities of the media as such, but are rather indications of the effectiveness of the materials which were developed for and presented in each medium. The reason for grouping and reporting results by media is to localize the variations in effectiveness of materials which may be attributable to teaching via different media. The results do not indicate comparisons of media made over identical content with identical test items, developed by the same writer, and employing identical presentation variables.

Table 12

Progress Check Mean Performance (Before Remediation) Across All Materials Developed in Each Medium, Second and Third Implementations

Media	Second Run	Third Run
CAI Script/IPB	87	96
Audiotape Script/IPB	93	95
CAI	88	94
Audiotape/IPB	93	94
Linear Text	88	94
Syndactic Text	87	92
Audiotape/Panelbook	82	89

The data for the third run in Table 2 was tabulated from the breakdown for media by segment given in Table 13. The breakdown within a segment in Table 13 indicates segments where a media option was available as was shown in Figure 2 (page 108).

It can be noted that the mean performance within each medium increased from the 2nd to the 3rd run, and that performance is quite comparable across all media in the final installation run.

Table 13

Mean Progress Check Performance (Before Remediation)

For Media By Segment for the Third Run

Audiotape/Panelbook (AT/PB), Syndactic Text (ST)			
Segment Number	Total Segment Performance	AT/PB	ST
2.2	90	91	90
2.3	85	84	85
2.4	87	88	87
2.5	94	91	95
Mean (2.2-2.5)	89	89	89
3.1	84	82	84
3.2	88	82	89
3.3	95	92	96
3.4	88	79	89
Mean (3.1-3.4)	89	84	89
5.7	95	85	96
5.8	93	93	92
5.9	90	88	90
5.10	94	95	94
Mean (5.7-5.10)	93	90	93
7.1	91	87	91
7.2	97	93	98
7.3	91	91	91
7.4	91	88	91
7.5	95	92	95
Mean (7.1-7.5)	93	90	93

Table 13 (Continued)

Segment Number	Total Segment Performance	AT/PB	ST
9.1	94	94	94
9.2	95	92	96
Mean (9.1-9.2)	95	93	95

Mean for all AT/PB = 89

Mean for all ST = 92

Syndactic Text

Segment No.	Segment Performance
1.1	91
2.1	91
2.6	86
2.7	81
2.8	91
2.9	91
3.5	93
5.1	88
5.2	85
5.3	93
6.1	91
6.2	93
6.3	92
6.4	95
8.4	95
8.5	93
8.6	94
11.1	93
11.2	97
11.3	92
Mean	91

Table 13 (Continued)

Audiotape/Intrinsically Programed Booklet (AT/IPB)
Audiotape Script/Intrinsically Programed Booklet (ATS/IPB)

Segment Number	Total Segment Performance	AT/IPB	ATS/IPB
4.4	96	95	96
4.5	94	93	95
4.6	93	93	93
4.7	93	91	94
Mean (4.4-4.7)	94	93	94
10.1	95	96	95
10.2	95	93	95
Mean (10.1-10.2)	95	95	95

Mean for all AT/IPB = 94

Mean for all ATS/IPB = 95

Computer Assisted Instruction (CAI)
CAI Script/Intrinsically Programed Booklet (CAIS/IPB)

Segment Number	Total Segment Performance	CAI	CAIS/IPB
12.1	97	96	97
12.2	94	95	94
12.3	96	94	97
12.4	94	92	95
Mean	95	94	96

Mean for all CAI = 94

Mean for all CAIS/IPB = 96

Table 13 (Continued)

Linear Text

Segment No.	Total Segment Performance
4.1	96
4.2	97
4.3	92
5.4	90
5.5	94
5.6	92
8.1	92
8.2	93
8.3	96

Mean for all Linear Texts = 94

VIII. SOME CONSIDERATIONS REGARDING THE DESIGN AND DEVELOPMENT FOR RESEARCH IN AN ON-GOING INSTRUCTIONAL SYSTEM

Introduction

Many studies of instructional variables have been conducted in a laboratory setting where precise control could be maintained. These types of studies might be termed basic research. The WLC research could also be termed basic but it was conducted in an applied setting. There are a multitude of factors to consider in establishing research in an ongoing course that will allow for control of extraneous variables that might affect the primary variables under investigation, and yet not hinder the learner's opportunity to learn nor sensitize him to the experimental manipulations. The factors considered in the WLC research effort are discussed in this chapter with the hope that it will be a useful guide for others considering research in an ongoing course system.

Experimental Design.

The experimental designs used to arrange the experimental instructional conditions in the research plan involved several common principles which were followed insofar as possible.

a. Conditions of instruction of primary experimental interest were always compared between alternate modules in the same CPT unit.

b. Students were randomly arranged in groups assigned to alternate modules in the CPT units. The primary test of achievement was the CPT,

which measured achievement over the entire unit under the conditions of a constant module.

c. Several widely separated CPT units involving the same medium were used in each experiment, with the same modules appearing in each unit.

d. Over the CPT units of a given experiment, each group of students experienced all types of modules involving the conditions of primary interest in one sequence of a counterbalanced set of sequences.

In technical terms, these principles may be summarized by the statement that Experiment I, II, and IV were designed as various types of repeated measure Latin Squares with CPT units defining the columns of the squares, randomized groups assigned to the rows of the squares, and modules corresponding to the counterbalanced latin letters of the squares. Experiment III was a mixed repeated-measure randomized block design, with repetition of modules and CPT units as blocks, and Experiment V was a completely randomized design (Meyers, 1967).

It is the purpose of this section to outline the basic problems of experimental control and course administration which lay behind the decisions to design the experimental conditions and sequences of events

according to these rules. It is believed a detailed consideration of the factors involved and the methods applied in the present case may provide considerable guidance in the design of research conducted in on-going course systems.

Design of the CPT unit. The CPT unit was so arranged as to accomplish several objectives relating to the effectiveness of the research and to course administration.

Two considerations were involved in the decision to use a group of segments as the basic research unit, rather than single segments. First, the unit of several successive segments more closely simulates the procedure under which the results of the research would be applied, where conditions found to be superior would be held constant over large sections, if not a whole course. Thus, it was desirable to provide conditions permitting the detection of delayed or slowly developing effects which might not appear immediately on the first administration of the conditions. Second, the single PC test was unlikely to provide a measuring instrument of sufficient sensitivity to demonstrate real efforts of the varied conditions of instruction. Limitations on the total student time which could be devoted to testing required that each PC be short, and given that effective

materials were prepared, there should be relatively little variation in immediate achievement on items directly referenced to segment objectives. Relatively low test reliabilities could therefore be expected both on the basis of test length and the criterion-referenced nature of the items.

The CPT was constructed to provide a norm-referenced test of greater reliability and discriminative power than could be achieved even by aggregating PC items across segments. Testing achievement of the unit as a whole following completion of the unit not only permitted testing of interrelationships and integration of the content of the segments at a higher level of complexity and difficulty, but provided measures at a point where some retention loss of learning in the early segments would have an opportunity to occur. The placement of the CPT at the end of a series of segments, then, could be expected to assess more accurately the amount of retainable learning achieved than would any testing conducted at the end of each segment. The contribution of retention loss to CPT performance would also be expected to increase sensitivity of the test to effects of the experimental conditions.

The decision to develop parallel modules within segments and to hold the medium constant over the unit was predicated on basic considerations of experimental

control. Segments of course content could be expected to vary considerably in difficulty and performance level. Additional variation would be contributed by variation in the effectiveness of materials developed to teach that content, and in the difficulty level of test items measuring achievement. The use of parallel modules with common content meant that conditions could be compared which varied only according to the presentation variables intended, with all other details of the presentation held precisely equivalent. In this fashion, most variation from sources associated with content was removed from the experimental comparisons into the columns of the Latin Squares, thereby enabling much smaller differences between conditions to be declared as statistically reliable.

Although the presentation variables can be regarded as fundamentally similar in different media, they cannot always be implemented according to exactly equivalent rules or criteria. The use of a single media in the CPT unit meant that the presentation variables could be manipulated in precisely the same way in each segment of the unit. Variation in the meaning and effects of the presentation variables was thus avoided, producing both a further gain in statistical precision, and a more explicit and unambiguous realization of the categories of the presentation variable, allowing easier interpretation of results.

It was, of course, necessary for groups of students to be assigned to a single module throughout the CPT unit so that their CPT performance would reflect the influence of only one constant set of experimental conditions. Since the student also changed media whenever beginning a unit, special instructions relating to the module could be conveniently incorporated in special instructions relating to the media without drawing undue attention to the variables being manipulated.

An additional consideration determining the plan of the CPT unit concerned the role of remediation. Remediation was required to bring all students to similar levels of competence following the measurement of experimental effects on the CPT. This procedure insured that no student would be disadvantaged by assignment to an inferior module except through loss of additional time for remedial learning. Furthermore, remediation insured that substantial differences in achievement among modules would not be carried over to the next research unit to confound the differences between modules compared there, nor to increase variation among students thus reducing the precision of the statistical evaluation.

The final point to be noted is the relation of the CPT unit to the final course system. The major features which demark the CPT unit are parallel modules and the CPT test. Since a single superior module may be selected and the alternatives abandoned, and the CPT may be eliminated, it should be clear that the CPT unit was readily adaptable to dismemberment in the ultimate use of the developed materials. The only trace of the CPT unit then remaining would be the points at which media change.

Media selection and sequencing. Selection of media for the instructional design of the USNA Leadership Management course was predicated on the requirements of the experimental designs, capacities needed for delivery of instructional presentations, and the diversity and flexibility expected of an individualized multimedia course. The media selected permitted precise experimental control and planned variaton in dimensions of stimulus representation, duration, response form, response demand frequency, and management decisions. Within limitations of existing facilities at USNA, media were selected which can be used in individually paced instruction without undue logistic difficulties, and with sufficient variety of instructional technique to maintain a consistent level

of student motivation. Some media were selected for their novel appeal, while the experimental manipulations of presentation design provided variety in the utilization of more traditional media forms.

Media placement and sequencing was limited to some extent by the number of segments required for each CPT unit and the number of segments in each part. Within these limitations, media were assigned to ensure perception of a sense of media variety, and to provide persistent media in segments with the most complex concepts.

The final media assignment to segments was carried out so as to provide a sufficient number of CPT units in the same media to accommodate the designed experiments, and to keep the CPT units of a given experiment widely separated in the course.

The purpose of having widely separated units in the same experiment with units assigned to other experiments intervening was to insure that any carryover effects of experimental conditions which remained after remediation were not carried systematically into the treatment conditions of the same experiment. Such carryover effects were, instead, randomized among the treatment conditions of different experiments. Thus, while the carryover effects might produce some increase in

variability and loss of precision in comparison of conditions of instruction, they were not allowed to systematically bias any comparison of conditions. Wide separation of CPT units in the same experiment also could be expected to systematically reduce problems of sequential correlation often associated with experiments of the repeated measure type.

The purpose of assigning the same medium to the CPT units of a particular experiment was the same as that given for holding the medium constant in the CPT unit, i.e., to permit the presentation variables to be manipulated in the same way and to produce similar effects in each segment of a given experiment. Since the presentation variables might possible have different effects in different media (media-presentation interactions), such variation in effects was avoided in all but Experiment III, where the interactions were of direct interest. Otherwise, the linear models for the statistical analyses would have been based on erroneous assumptions, and the interpretation of results would be somewhat difficult.

Arrangement of systems and modules. Several advantages accompanied the use of counter-balanced sequences of modules across the majority of CPT units. First, each student encountered all the major module variations, roughly equating exposure to relatively good or poor conditions of instructions. In addition to remediation, the equation of experimental histories produced by counterbalancing ensured that no student was handicapped in opportunity to obtain good grades through consistent assignment to inferior conditions. This was a distinct administrative advantage in the assignment of grades, since no special correction or subdivision of students was required to account for differential effects of experimental history.

Measurement of performance of each student on each type of module permitted the evaluation of the primary conditions of interest on a within-student basis. Within-student comparisons involve a marked gain in precision (reduction in random variability) since the variation among students does not contribute to differences between experimental conditions. The gain in precision produced by this means was especially important because of the small number of students available.

Experience with within-student designs also indicates

that the treatment differences found in such designs tend to be more characteristic of all individual students and less fragile in the face of alterations in the context and preceding events. Only relatively consistent and durable experimental differences remain after averaging over such diversity of content, and module and media histories as are produced by the counterbalanced variations.

With the gain in precision produced by counterbalancing, further gains from matching groups or statistical control through analysis of covariance were not deemed necessary, permitting random assignment of students to groups. There was no assurance that any variable available for matching or analysis of covariance would have a sufficient relationship to CPT performance to produce any substantial gain in precision. Furthermore, the use of either of these procedures would have increased the complexity of the design and/or statistical analysis to the point of unmanageability. Corrections for student withdrawals and missing data, as discussed below, would have been much more difficult as well.

Two other considerations also favored the use of randomized groups. Evaluation (with lesser precision) of certain media and presentation main effects and interactions of secondary interest was permitted on a between-

student basis through judicious arrangement of the Latin Squares. Furthermore, randomization provided that carryover effects, as discussed previously, were not allowed to systematically enter the comparison of conditions of instruction.

Analysis of Variance Methods

Four sets of data on each experiment were analyzed, including PC data totaled over the segments of each CPT unit, CPT total scores, CPT Type I scores, and CPT Type II scores. Wherever the test scores were based on differing number of test items for separate CPT units, the original scores were concerted to percentages prior to analysis.

Analysis of variance was performed on each set of data based on standard linear model methods for the types of designs involved (see especially Winer, 1962, Ch. 7, 8, and 10, and Meyers, 1966, Ch. 8, 9, and 10), with some modifications required as described below. Although multivariate analysis was jointly applicable to the different measures obtained, the univariate analyses were preferred for ease of computation and interpretation in the light of the complexity of the designs.

Inspections of residuals indicated reasonable satisfaction of the required statistical assumptions, so no statistical tests of these assumptions were performed.

The small number of conditions compared in any one experiment also obviated the need for multiple-comparison procedures to aid in the interpretation of results found significant by overall F-tests.

Two problems did arise, however, which required special techniques of analysis. First, there were several cases in which individual CPT scores were inaccurate resulting from minor errors in implementation of experimental procedures. Such scores were dropped from the analysis and replaced by least squares estimates. The estimates were obtained by following the procedure developed by Yates for the estimation of missing data in randomized block designs as given in Cochran and Cox (1957, p. 110). This procedure was appropriate since the data of a single group within a repeated-measure Latin Square design forms a randomized-block design when students are identified as blocks. Computation was based on the two-way student x unit table from which the score was dropped.

The second problem was that in both research runs some students withdrew from the course, resulting in unequal group sizes in all experiments. Since the loss of students was unrelated to the nature of the experimental conditions, and the group sizes which remained were not very disparate, the computational procedure for unweighted-means analysis

of variance was followed in all cases (Winer, 1962, p. 374-378). In this form of analysis, components of error variance are estimated from the original data of each group, but the analysis of treatment effects is based on tables of unweighted means. The use of unweighted means causes all experimental conditions to contribute equally to the estimation of effects, without regard to the number of individuals in those conditions.

REFERENCES

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives: The classification of educational goals. (Handbook I: Cognitive Domain) New York: Longmans, Green and Co., 1956.

Briggs, Leslie J. Handbook of Procedures for the Design of Instruction. American Institutes for Research. Pittsburgh, Pennsylvania, 1970. pp. 1-206.

Brudner, H. J. Computer-managed instruction--conclusion. Instructional Media, Vol. 1, No. 2, May, 1969.

Campeau, Peggie L. Level of anxiety and presence or absence of feedback in programed instruction. Palo Alto, California: American Institutes for Research, February, 1966. (Mimeo.)

C'de Baca, J., and Chadwick, C. B. PROMOD: A new approach to multi-level programming. NSPI Journal, 1968, 7 (2), pp. 8-10.

Cochran, W. D., and Cox, G.M. Experimental Designs. Second Edition, New York, Wiley, 1957.

Cronbach, Lee J., Chapter 2: How Can Instruction be Adapted to Individual Differences? in Robert M. Gagne, Learning and Individual Differences, Charles E. Merrill Books, Inc., Columbus, Ohio, 1967, pp. 1-265.

Draper, N. R. and Smith, H. Applied Regression Analysis. New York: Wiley, 1966.

(Continued on next page)

Ellis, H. C. Research design in programmed learning. National Society for Programmed Instruction. August 1962, pp. 7-8.

Evans, J. L., Homme, L. E., & Glaser, R. The RULEG system for the construction of programed verbal learning sequences. Journal of Education Research, 1962, 55, pp. 513-518.

Flanagan, J. C. Functional education for the 70's. Phi Delta Kappan, 49 (Sept. 1967), pp. 27-32.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Holland, J. G. Research on programming variables. In Teaching Machines and Programmed Learning II (Glaser, R., Ed., National Education Association, Washington, D. C., 1965).

Mager, Robert F. Preparing Instructional Objectives. Fearon Publishers, Palo Alto, California. 1962.

Meyers, D. E. Adjunct to self-study for aircrew refresher training under operational conditions in the Air Defense Command. Dayton, Ohio: Wright-Patterson Air Force Base, Behavioral Sciences Laboratory, 1965. (AMRL-TR-65-83).

Reid, J. C. and MacLennan, P. W. Research in Instructional Television and Film. U. S. Department of Health, Education, and Welfare, Bureau of Research, Catalog #FS-5.234:34041, 1967.

(Continued on next page)

Stolurow, L. M. Implications of current research and future trends. Journal of Educational Research, 1962, 55, pp. 518-27.

Tosti, D. T., and Ball, J. R. A Behavioral Approach to Instructional Design and Media Selection. Albuquerque, 1969. Behavior Systems Division, Westinghouse Learning Corporation.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

APPENDIX A

PREDICTOR VARIABLES

Variable Code:			Variable Name:
STV	1	*	SAT - verbal
STM	2	*	SAT - math
ENC	3	*	CEEB English Comprehension
MAT	4	*	CEEB Math Achievement
RNK	5	*	Converted rank in class
REC	6		Recommendation score
GPA	7	*	Grade point average
EDWARDS PERSONAL PREFERENCE SCHEDULE (EPPS)			
ACV	8	*	Achievement
DEF	9		Deference
ORD	10	*	Order
EXH	11		Exhibition
AUT	12	*	Autonomy
AFF	13		Affiliation
ISP	14		Intraspection
SUC	15		Succorance
DOM	16		Dominance
ABA	17		Abasement
NUR	18		Nurturance
CHG	19		Change
END	20		Endurance

EDWARDS PERSONAL PREFERENCE SCHEDULE (Cont'd)

HET	21			Heterosexuality
AGG	22			Aggression
VAL	23			Validity Scale

16 PERSONALITY FACTOR SCALE (16PF)

PFA	24		A	Reserved vs Outgoing
PFB	25	*	B	Concrete thinking vs Abstract thinking
PFC	26		C	Affected by feelings vs Emotionally stable
PFE	27		E	Humble vs Assertive
PFF	28		F	Sober vs Happy-go-lucky
PFG	29		G	Expedient vs Conscientious
PFG	30		H	Shy vs Venturesome
PFI	31		I	Tough minded vs Tender minded
PFL	32		L	Trusting vs Suspicious
PFM	33		M	Practical vs Imaginative
PFN	34		N	Forthright vs Shrewd
PFO	35		O	Placid vs Apprehensive
PF1	36		Q_1	Conservative vs Experimenting
PF2	37		Q_2	Group-dependent vs Self sufficient
PF3	38		Q_3	Undisciplined self-conflict vs Controlled
PF4	39		Q_4	Relaxed vs Tense

16 PERSONALITY FACTOR SCALE (SECOND ORDER FACTORS)

EXT	40		Extraversion
ANX	41		Anxiety
TOP	42		Tough Poise
IND	43		Independence
NEU	44		Neuroticism
LEA	45		Leadership
CRE	46		Creativity

OHIO STATE PSYCHOLOGICAL (OSU)

OS1	47	*	Test 1 Same-Opposite Section
OS2	48	*	Test 2 Analogy Section
OS3	49	*	Test 3 Reading Comprehension Section
OS4	50	*	Test 4 Total Reading

STRONG VOCATIONAL INTEREST BLANK (SVIB)

NAV	51	Naval Officer
PTH	52	Physical Therapist
DEN	53	Dentist
OST	54	Osteopath
VET	55	Veterinarian
DOC	56	Physician
PYI	57	Psychiatrist
PYO	58	Psychologist
BIO	59	Biologist
ARC	60	Architect

STRONG VOCATIONAL INTEREST BLANK

MTH	61	Mathematician
PHY	62	Physicist
CHE	63	Chemist
ENG	64	Engineer
PMR	65	Production Manager
ARM	66	Army Officer
AFO	67	Air Force Officer
CAR	68	Carpenter
FOR	69	Forest Service Man
FAR	70	Farmer
MST	71	Math-Science Teacher
PRI	72	Printer
POL	73	Policeman
PDR	74	Personnel Director
PAD	75	Public Administrator
RCO	76	Rehabilitation Counselor
YMS	77	YMCA Secretary
CRA	78	Community Recreation Admin.
SWO	79	Social Worker
SSC	80	Social Science Teacher
SSU	81	School Superintendent
MIN	82	Minister
LIB	83	Librarian

STRONG VOCATIONAL INTEREST BLANK (Cont'd)

ART	84	Artist
MUP	85	Musician Performer
MUT	86	Music Teacher
CPO	87	CPA Owner
CPA	88	Senior CPA
ACC	89	Accountant
OWO	90	Office Worker
CMR	91	Credit Manager
COC	92	Chamber of Commerce Exec.
BET	93	Business Education Teacher
PUR	94	Purchasing Agent
BAN	95	Banker
PHA	96	Pharmacist
MOR	97	Mortician
SMR	98	Sales Manager
RES	99	Real Estate Salesman
INS	100	Life Insurance Salesman
ADV	101	Advertising Man
ATY	102	Lawyer
AUT	103	Author-Journalist
PMF	104	President, Mfg. Concern
CPR	105	Computer Programmer
INT	106	Interpreter (language)

STRONG VOCATIONAL INTEREST BLANK (Cont'd)			
A-B	107		Therapists (with Schizophrenics)
ACH	108	*	Academic Achievement
L-C	109		Confidential scale relating to predicted job tenure
M-F	110		Masculinity-Femininity
OCL	111		Occupational Level
SIN(OIE)	112		Occupational Introversion-Extroversion
SPL	113		Specialization Level
N-6	114		NROTC Officer (predicted tenure)
MGE(MO)	115		Managerial Orientation
STUDENT QUESTIONNAIRE - HIGH SCHOOL OR COLLEGE SUBJECTS STUDIED (SQ)			
S01	116		Psychology
S02	117		Sociology
S03	118		Business
S04	119		Human Relations (or equivalent)
S05	120		Leadership
METHODS OF PREVIOUS INSTRUCTION (SQ)			
S06	121		Team teaching
S07	122		Computer-aided instruction
S08	123		Teaching machine
S09	124		Programed textbook
S10	125		Television
S11	126		Videotape
S12	127		Audiotape

STUDY HABITS (SQ)		
S13	128	Rate study habits
S14	129	Average hours of study
S15	130	Anticipated hours studying Leadership
COLLEGE RELATED ABILITIES (SQ)		
S16	131	General college achievement
S17	132	Vocabulary
S18	133	Reading ability
S19	134	Writing ability
S20	135	Oral expression
S21	136	Verbal participation in class
S22	137	Pace in classroom activities

* One of the 15 preselected predictor variables

UNITED STATES NAVAL ACADEMY
LEADERSHIP MANAGEMENT COURSE

STUDENT QUESTIONNAIRE

NAME (print)_____
(Last) (First) (Middle)

ALPHA CODE ☐ ☐ ☐ ☐ ☐ ☐

CLASS

☐ 1. 1970
☐ 2. 1971
☐ 3. 1972
☐ 4. 1973

HIGH SCHOOL OR COLLEGE SUBJECTS STUDIED

1. Psychology
   ☐ 1. less than one semester
   ☐ 2. one semester
   ☐ 3. two semesters
   ☐ 4. more than two semesters

2. Sociology
   ☐ 1. less than one semester
   ☐ 2. one semester
   ☐ 3. two semesters
   ☐ 4. more than two semesters

3. Business
   ☐ 1. less than one semester
   ☐ 2. one semester
   ☐ 3. two semesters
   ☐ 4. more than two semesters

4. Human Relations (or equivalent)
   ☐ 1. less than one semester
   ☐ 2. one semester
   ☐ 3. two semesters
   ☐ 4. more than two semesters

HIGH SCHOOL OR COLLEGE SUBJECTS STUDIED, continued

5. Leadership

- ☐ 1. less than one semester
- ☐ 2. one semester
- ☐ 3. two semesters
- ☐ 4. more than two semesters

METHODS OF INSTRUCTION BY WHICH YOU HAVE BEEN TAUGHT

6. Team Teaching

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

7. Computer-Aided Instruction

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

8. Teaching Machine

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

9. Programed Textbook

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

10. Television

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

METHODS OF INSTRUCTION BY WHICH YOU HAVE BEEN TAUGHT, continued

11. Videotape

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

12. Audiotape (tape recorder)

- ☐ 1. none
- ☐ 2. less than 3 weeks
- ☐ 3. 3 to 6 weeks
- ☐ 4. 6 to 12 weeks
- ☐ 5. more than 12 weeks

STUDY HABITS

13. Would you rate your study habits

- ☐ 1. poor
- ☐ 2. fair
- ☐ 3. good
- ☐ 4. very good

14. On the average, do you study

- ☐ 1. less than 6 hours a week
- ☐ 2. 6 to 10 hours a week
- ☐ 3. 10 to 14 hours a week
- ☐ 4. more than 14 hours a week

15. Approximately how much time do you anticipate studying leadership per week (including class time)?

- ☐ 1. less than 4 hours
- ☐ 2. 4 to 6 hours
- ☐ 3. 6 to 8 hours
- ☐ 4. more than 8 hours

COLLEGE-RELATED ABILITIES

16. General College Achievement

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

COLLEGE-RELATED ABILITIES, continued

17. Vocabulary

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

18. Reading Ability

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

19. Writing Ability

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

20. Oral Expression

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

21. Willingness to participate verbally in class

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average

22. Ability to keep pace in classroom activities

- ☐ 1. very much below average
- ☐ 2. below average
- ☐ 3. average
- ☐ 4. above average
- ☐ 5. very much above average